

ScholarWorks@GSU

Essays on Fiscal Policy and Tax Compliance

Authors	Small, Oronde D
DOI	https://doi.org/10.57709/10276603
Download date	2026-04-21 07:04:45
Link to Item	https://hdl.handle.net/20.500.14694/4417

8-31-2017

Essays on Fiscal Policy and Tax Compliance

Oronde D. Small

ABSTRACT
ESSAYS ON FISCAL POLICY AND TAX COMPLIANCE

By
ORONDE DIA SMALL

August 2017

Committee Chair: Dr. Sally Wallace

Major Department: Economics

This dissertation comprises three essays that examine critical aspects of fiscal policy and explores important determinants of tax compliance in a developing country context. The first essay examines the fiscal response to changes in debt-to-GDP for a panel of developing countries. Our empirical strategy adopts a dynamic difference generalized methods of moments (DGMM) model with forward orthogonal deviation. We find a positive and significant response for the primary balance and ‘fiscal effort’ to changes in debt-to-GDP. For the fiscal components, we find a positive relationship between debt-to-GDP and general and tax revenues, and a negative relationship with primary spending. We also find evidence of nonlinearities, with countries making larger increases in the primary balance and fiscal effort at higher levels of debt, largely driven by increases in revenues. Higher income countries demonstrate a greater propensity to adjust along the revenue margins, compared to lower income countries. This might be indicative of systemic revenue mobilization challenges facing the latter.

The second essay examines the effect of the provision of taxpayer services on filing and payment of the corporate income tax (CIT) and general consumption tax (GCT) for large taxpayers in Jamaica. We use a regression discontinuity design (RDD) that exploits an exogenous jump in the intensity of taxpayer service delivery, which occurs when a taxpayer

reaches gross receipts of J\$500 million (US\$5.7 million) and is selected into the large taxpayer office (LTO). The results indicate null effects for the CIT but positive filing and payment compliance effects for the GCT. The contrasting results for the CIT and GCT may be due to the relatively weaker legal enforcement framework of the former. The results provide suggestive evidence of a complementarity between the strength of the legal enforcement framework of the taxing regime and the provision of taxpayer services.

In the third essay we implement public goods messaging experiments to examine the effects on personal income tax (PIT) compliance among self-employed individuals in Jamaica. In the first sub-experiment we examine the effect of the standard public goods message on payment of quarterly PIT obligations. In the second sub-experiment we focus on payment of PIT arrears and expand the message context to include a variant of the standard public goods message – which provides additional information on actual spending on key public goods and services. The compliance outcomes in sub-experiment two relate to established PIT delinquencies, compared to sub-experiment one where there is no legal obligation to comply. We find that the standard public goods message had no effect on compliance with quarterly PIT payments in sub-experiment one. However results from sub-experiment two indicate positive compliance effects from the standard and augmented public goods messages on the probability of making a payment and the amount of PIT arrears paid after nineteen weeks. Point estimates from the standard and augmented public goods messages are not statistically different for any of the outcomes examined; suggesting that additional information on public spending allocations does not matter.

ESSAYS ON FISCAL POLICY AND TAX COMPLIANCE

By

ORONDE DIA SMALL

A Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree

of

Doctor of Philosophy

in the

Andrew Young School of Policy Studies

of

Georgia State University

GEORGIA STATE UNIVERSITY

2017

Copyright
Oronde Dia Small
2017

ACCEPTANCE

This dissertation was prepared under the direction of Oronde Small's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair:

Dr. Sally Wallace

Committee:

Dr. Nguedia Pierre Nguimkeu

Dr. Felix K. Rioja

Dr. Neven T. Valev

Electronic Version Approved:

Sally Wallace, Dean
Andrew Young School of Policy Studies
Georgia State University
August, 2017

ACKNOWLEDGMENT

I thank the members of my dissertation committee and in particular Dr. Sally Wallace, for their guidance and support. I would also like to thank my family for their unwavering support throughout the process. Finally to my wife Kandise Small and son Zuhri Small, thank you for your sacrifice, support and constant encouragement, without which I could not have completed this dissertation.

TABLE OF CONTENTS

ACKNOWLEDGMENT.....	iv
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
INTRODUCTION.....	1
Essay 1: An Examination of the Fiscal Policy Response to Public Debt in Developing Countries	7
Introduction.....	7
Fiscal Policy in Developing Countries.....	9
Theoretical Framework.....	10
Literature Review.....	13
Data.....	16
Empirical Strategy.....	19
Results.....	21
Results for the Full Sample.....	21
Non Linear Fiscal Reaction.....	26
Heterogeneous Effects.....	30
Sensitivity Analysis.....	34
Conclusion.....	35
Essay 2: The Effect of Taxpayer Service Provision on Tax Compliance for Large Taxpayers in Jamaica.....	38
Introduction.....	38
Background.....	42
The Corporate Income Tax and the General Consumption Tax.....	42
The Large Taxpayer Office.....	43
Theory.....	44
Data and Empirical Strategy.....	47
Data.....	47
Empirical Analysis.....	50
Results.....	54
Filing Compliance for the CIT.....	54
Payment Compliance for the CIT.....	56
Filing Compliance for the GCT.....	58

Payment Compliance for the GCT	59
Heterogeneous Effects	63
Conclusion.....	64
Essay 3: The Effect of Public Goods Messaging on Personal Income Tax Compliance among the Self-Employed: Evidence from a Field Experiment in Jamaica.....	66
Introduction	66
Background	68
Literature Review	70
Data and Experimental Design.....	73
Data.....	73
Experiment Design	74
Sub-experiment One: Effect of Public Goods Messaging on Third and Fourth Quarter PIT Payments.....	75
Sub-experiment Two: Effect of Public Goods Messaging on Payment of Personal Income Tax Arrears	77
Randomization Balance.....	78
Empirical Strategy	80
Results	81
Sub-experiment One	81
Sub-experiment Two	84
Effect of Treatment on the Treated (TOT) for Sub-experiment One and Two	89
Robustness and Heterogeneity.....	91
Cost Benefit Analysis.....	96
Conclusion.....	97
CONCLUSION.....	100
Appendix A: Appendix for Essay One	103
Appendix B: Appendix for Essay Two.....	114
Appendix C: Appendix for Essay Three.....	122
References.....	128
VITA	135

LIST OF TABLES

Table 1: Summary Statistics, 1990 – 2011	17
Table 2: Results for Primary Balance	22
Table 3: Results for the Fiscal Effort	23
Table 4: Results for Fiscal Components	25
Table 5: Results for Non Linear Fiscal Response	29
Table 6: Results for the Fiscal Response for Low and Lower Middle Income Countries	32
Table 7: Results for the Fiscal Response for High and Upper Middle Income Countries	33
Table 8: Summary Statistics, CIT & GCT	49
Table 9: Effect of the LTO on Timely Filing of the CIT	55
Table 10: Effect of the LTO on Payment Compliance for the CIT	57
Table 11: Effect of the LTO on Timely Filing of the GCT	59
Table 12: Effect of the LTO on Payment Compliance for the GCT	61
Table 13: Summary Statistics Sub experiment One and Two	74
Table 14: Covariate Balance for Sub-experiment One	79
Table 15: Covariate Balance for Sub-experiment Two	80
Table 16: Probability of Making Quarterly PIT Payments	83
Table 17: Amount of Quarterly PIT Paid	84
Table 18: Effect of Public Goods Messages on Probability of Payment of PIT Arrears	87
Table 19: Effect of Public Goods Messages on Payment of PIT Arrears	88
Table 20: TOT Effect of Public Goods Message on Quarterly PIT Payment	90
Table 21: TOT Effect of Public Goods Message on Payment of PIT Arrears	91
Table 22: Placebo Test – Payment of PIT Arrears Six Weeks Before and After Experiment	93
Table 23: Impact of Public Goods Messages on Payment of PIT Arrears by Age Cohort	95
Table 24: Effect of Public Goods Messages on Payment of PIT Arrears, by Level of PIT Arrears	96

LIST OF FIGURES

Figure 1: Debt-to-GDP Ratio, 1990 – 2011	18
Figure 2: Test of Identification - First Stage RDD	52
Figure 3: Effect of the LTO on Timely Filing of the CIT	54
Figure 4: Effect of the LTO on Timely Payment of the CIT	56
Figure 5: Effect of the LTO on Timely Filing of the GCT	58
Figure 6: Effect of the LTO on Payment Compliance for the GCT	60
Figure 7: PIT Payment Rates by Quarter	82
Figure 8: Average PIT Payment by Quarter	82
Figure 9: Payment Rate for PIT Arrears	85
Figure 10: Cumulative Payment for PIT Arrears per Person.....	86
Figure 11: Coefficient Estimates of Treatment Effects over Time.....	92

INTRODUCTION

A primary focus of fiscal policy in developing countries centers on debt management. Reinhart, Rogoff and Savastano (2003) point out that the ‘debt intolerance’ that characterizes many developing countries implies a heightened sensitivity on the part of creditors, which ultimately limits access to credit markets. Access to credit markets is therefore an important determinant of fiscal policy in many developing countries. This argument is supported by Albero and Montero (2006) who maintain that public debt sustainability accounts for the whole pro-cyclicality of fiscal policy in Latin America. This is evidenced by countries commitment to fiscal consolidation, even during harsh economic times, perhaps as a signal to creditors that debt remains a priority. However, notwithstanding the impetus to react, developing countries may be limited in the extent to which they are able to do so. These limitations may act on countries overall ability to increase primary surpluses in response to growing debt, and can possibly be more of a binding constraint on either the revenue or expenditure side of the budget.

Understanding the anatomy of the fiscal policy response to debt in developing countries is an interesting but largely unexplored question. Existing evidence point to significant revenue mobilization challenges across developing countries, which might compromise their ability to adjust along the revenue margin. If countries are so constrained, then strategies to unlock the blockages to revenue growth are an obvious priority. A low level of tax compliance has been identified as one impediment to revenue growth across developing countries. In this context understanding what are some of the key factors that drive tax compliance in these countries becomes particularly important.

This dissertation comprises three essays that examine critical issues in fiscal policy and tax compliance in developing countries. The first essay examines the relationship between the primary balance (revenues less primary expenditure), fiscal effort (tax revenues less primary expenditure) and their respective fiscal components – general revenue, tax revenue and primary expenditure – with the lagged value of the debt-to-GDP ratio for a panel of developing countries. We contribute to the existing literature in several important ways. Firstly, although previous studies examine the fiscal response of the primary balance to changes in debt-to-GDP for advanced countries, scant attention has been paid to developing countries. Secondly, we go further than most previous studies and examine specific margins of response. In particular we examine the response of countries ‘fiscal effort’, defined here as total consolidated central government tax revenues less non-interest expenditures. Relative to the primary balance, ‘fiscal effort’ does not include grants or other non-tax revenues, and arguably captures a fiscal measure that is more closely linked to government fiscal policy discretion. Further we dichotomize the fiscal reaction function (FRF) into the general revenue, tax revenue and primary spending components and examine each separately. Whereas several studies examine the FRF, very few have examined the reaction of these separate components and none have done so for developing countries. We also test for nonlinearities across the debt distribution and heterogeneity in the fiscal policy response between low / lower-middle income and high / upper-middle income countries.

Theoretical models of debt sustainability predict a positive relationship between primary balance and debt. It follows, that the expected relationship between debt-to-GDP and the revenue components is positive, and that with primary spending is negative. We test the theoretical predictions using data for a panel of 54 developing countries over the period 1990 – 2011. Our

empirical strategy adopts a dynamic difference generalized methods of moments (DGMM) model with forward orthogonal deviation (FOD) (Arellano & Bover, 1995). We find a positive and significant response for the primary balance and ‘fiscal effort’ to changes in debt-to-GDP. For the full sample we find evidence that countries adjust along both the revenue and expenditure margins at roughly the same rate. This sort of measured adjustment might be indicative of deliberate policy on the part of governments, not to engage in asymmetric policy changes that could impact negatively on overall economic welfare. We also find suggestive evidence on nonlinearities in fiscal response, with countries making larger increases in the primary balance and fiscal effort at higher levels of debt, mainly driven by increases in revenues. At lower levels of debt countries tend to use spending cuts. Tests for heterogeneity suggests that high / upper-middle income countries have a greater propensity to adjust along the revenue and tax revenue margins compared to low / lower middle income countries. This might be indicative of systemic revenue mobilization challenges facing the latter.

The second essay complements the first and investigates potential factors that impinge on the revenue mobilization capacity of developing countries. In particular we examine the effect of the provision of taxpayer services on filing and payment of the corporate income tax (CIT) and general consumption tax (GCT) for large taxpayers in Jamaica. Despite the huge dependence of developing country governments on large taxpayers for almost all their revenue, very little has been done to study the effects of tax administration interventions on compliance for this taxpayer segment. Further, the emergence of complementary theories of tax compliance and the adoption of administration interventions that seek to leverage non-pecuniary factors have spawned a literature that examine the effects of these factors on tax compliance in developing countries. However, very few focus on the provision of taxpayer services. This essay attempts to fill these

gaps in the literature. Another important contribution of this essay is that we explore the effects of taxpayer services conditional on the strength of the legal enforcement framework of the taxing regime, to test the relationship between service delivery and enforcement strength.

We focus on the taxpayer's decision to file and pay taxes conditional on their reporting decision. These are important margins of response particularly in developing countries with relatively weak tax administrations that find it difficult to collect reported taxes. The empirical strategy adopts a regression discontinuity design (RDD) that exploits an exogenous jump in the intensity of taxpayer service delivery, which occurs when a taxpayer reaches gross receipts of J\$500 million (US\$5.7 million) and is selected into the large taxpayer office (LTO). This approach compares the compliance behavior of those located just to the right of the threshold – who are selected into the LTO, to otherwise similar taxpayers located just the left of the threshold – who are marginally not selected into the LTO. Assuming all other key taxpayer characteristics transition smoothly across the threshold, the RDD estimates are causal. The results indicate null effects for the CIT but positive filing and payment compliance effects for the GCT. The contrasting results for the CIT and GCT may be due to the relatively weaker legal enforcement framework of the former. The results provide suggestive evidence of a complementarity between the strength of the legal enforcement framework of the taxing regime and the provision of taxpayer services.

In the third essay we continue to explore the effect of non-pecuniary factors on tax compliance in Jamaica, by examining the role of public goods messages. Compared to the second essay, which focuses on large taxpayers and on the CIT and the GCT, this essay focuses on the self-employed who are typically small to medium sized, and on the personal income tax (PIT). The contribution of this paper is three fold. Firstly we add to a nascent but growing

literature that uses messaging experiments to study casual determinants of tax compliance in developing countries (Castro & Scartascini, 2013; Ortega & Sanguinetti, 2013; Pomeranz, 2013; Carillo, Pomeranz & Singhal, 2014; Brockmeyer, Hernandez & Kettle, 2016). This is the first in an English speaking Latin American Caribbean (LAC) country. Secondly we examine the compliance effects of taxpayer knowledge about spending allocations across important public goods and services – a proxy for taxpayer agency. Whereas recent studies have examined similar questions in a controlled lab environment, we move the analysis into the field. Thirdly the compliance outcomes we examine offers an opportunity to compare the effect of public goods messages (and non-pecuniary factors more generally) on compliance with a purely voluntary obligation – for which there are no penalties for non-compliance, and obligations for which non-compliance attract significant fines.

We implement a randomized messaging experiment to examine the effect of public goods messages on PIT compliance amongst self-employed individuals in Jamaica. We conduct two sub-experiments and examine the effect of the messages on two sets of compliance outcomes. In the first sub-experiment we examine the effect of the standard public goods message on payment of quarterly PIT obligations. Because there are no penalties for non-payment of quarterly PIT, the compliance effects will be free of confounding from enforcement action – a truly zero enforcement environment. In this context, the structure of the PIT penalty regime in Jamaica provides a unique opportunity to examine the effect of the standard public goods message on voluntary compliance. In the second sub-experiment we focus on payment of PIT arrears and expand the message context to include an augmented public goods message, which provides additional information on actual spending on key public goods and services. Providing information on spending allocation can ‘open up the books’ and can elicit trust and arguably

provide a sense of agency in the mind of taxpayers. In this sub-experiment the compliance outcomes relate to established PIT delinquencies, compared to sub-experiment one where there is no legal obligation to comply.

We find that the standard public goods message had no effect on compliance with quarterly PIT payments in sub-experiment one. However results from sub-experiment two indicate positive compliance effects for the probability of making a payment and the amount of PIT arrears paid, after nineteen weeks. Point estimates from the standard and augmented public goods messages are not statistically different for any of the outcomes examined; suggesting that additional information on public spending allocations does not matter. Tests for heterogeneity in compliance response suggest that compliance gains are driven largely by older taxpayers. Unlike previous studies we find only limited evidence of differential compliance response between individuals with high and low tax debt.

The remainder of the dissertation provides a more detailed discussion of each essay.

Essay 1: An Examination of the Fiscal Policy Response to Public Debt in Developing Countries

Introduction

Debt sustainability has long been a major concern for governments in both developed and developing countries. However government's capacities to sustain their fiscal accounts differ tremendously across the two groupings. There is a large but still emergent literature devoted to examining the fiscal reaction functions (FRF) for developed countries. However, much less focus has been given to this question in the context of developing countries. This is surprising especially since developing countries are likely to face less favorable capital market access conditions and are thus much more exposed to fiscal sustainability challenges (Alberola & Montero, 2006).¹ Unlike their more developed counterparts, public debt in developing countries has more immediate consequences for the economy. As a group they share a colored history of debt crises and subsequent periods of painful economic adjustments. The recurrences of debt crises invite a closer enquiry into whether fiscal policies in these countries have been deployed responsibly.

This essay estimates the FRF for a panel of developing countries. The FRF captures changes in the primary balance, defined as consolidated central government revenues less non interest expenditures, in response to a change in the one period lag debt-to-GDP ratio, after controlling for *inter alia* volatility in economic output and government spending. Additionally, we go further than most previous studies and examine specific margins of response. In particular we examine the response of countries 'fiscal effort', defined here as total consolidated central government tax revenues less non-interest expenditures. Relative to the primary balance, 'fiscal

¹ Developing countries are generally more susceptible to natural disasters and other economic, social and political disturbances that could compromise their ability to remain solvent.

effort' does not include grants or other non-tax revenues, and arguably captures a fiscal measure that is more closely linked to government fiscal policy discretion.² Further we dichotomize the FRF into the general revenue, tax revenue and primary spending components, and examine each separately. Whereas several studies examine the FRF, very few have examined the reaction of these separate components and none have done so for developing countries. This is an important contribution since it provides a test of how the different sides of the budget are able to adjust in response to changes in public debt. Moreover, the estimated elasticities with respect to the debt-to-GDP ratio can be used to inform appropriate fiscal policy adjustments for the countries examined. We also test for nonlinearities in the fiscal policy response conditional on the level of debt, and examine heterogeneous responses across low / lower-middle income and high / upper-middle income countries.

The results indicate a positive response in the primary balance and 'fiscal effort' to changes in debt-to-GDP. In general, countries in the full sample appear to adjust along both the revenue and expenditure margins. We find some evidence of nonlinearities in the fiscal response – countries make larger adjustments in the primary balance and 'fiscal effort' at relatively higher levels of debt. Interestingly, fiscal adjustments at higher debt levels are driven largely by increases in revenues, whereas the tendency is to cut primary spending when debt is relatively low. We find that low / lower-middle and high / upper-middle income countries have been fiscally responsible – with a generally positive FRF. However compared to low / lower-middle income countries, high / upper-middle income countries demonstrate a greater propensity to adjust along the general revenue and tax revenue margins. This might be indicative of systemic revenue mobilization challenges facing the latter.

² The 'fiscal effort' measure places a stricter requirement on governments to use internally generated tax revenues to sustain debt.

Fiscal Policy in Developing Countries

There is some legitimacy for government's use of deficit financing to fund critical expenditures. The classic Keynesian argument advances that deficit financing to boost aggregate demand and employment is not only justified but necessary. This argument is in part based on the assumption that there are no intergenerational equity concerns associated with debt accumulation, as the burden of public debt is assumed to be borne by the generation that issued the debt.³ A related argument comes from Barro (1979) who advances a theory of tax smoothing as justification for the use of deficit financing to smooth out cyclical fluctuations in output. In this context, deficit financing, used as a substitute for distortionary adjustments in tax rates can be welfare improving.

In theory a country can afford to incur debt if, over time it is able to consistently generate higher relative growth rates in real economic output or where it is able to consistently generate increases in primary surpluses that exceed the interest rate growth rate differential. However, there is a limit on the extent to which a country is able to continue to accumulate debt before it begins to raise concerns about the sustainability of its fiscal policy. There is some evidence to suggest that this debt limit is much lower in developing countries relative to more advanced economies. Reinhart, Rogoff and Savastano (2003) highlight the "debt intolerance" that characterizes many developing countries, which heightens sensitivity on the part of creditors who may in turn limit access to credit markets. In the extreme case, effective marginal interest rates on debt can become infinite and countries can essentially be shut out of the credit market altogether (Flood & Marion, 2006; Gosh, Kim, Mendoza, Ostry & Qureshi, 2011). Access to

³ This argument has come up against some criticism. Barro (1974) outlined that the issuing of debt within the context of a deficit financing program geared toward increasing government expenditures and boosting aggregate demand may not have the desired result. He argues that the future taxation implied by the accumulation of debt will have offsetting effects on household income and on their propensity to consume. Moreover there are costs associated with issuing debt that if incorporated could result in negative overall welfare effects for households.

credit markets is therefore an important determinant of fiscal policy in developing countries.⁴

This argument is supported by Alberro and Montero (2006) who argue that public debt sustainability accounts for the whole pro-cyclical fiscal policy in Latin America. However, notwithstanding the impetus to react, developing countries may be limited in the extent to which they are able to do so.

Theoretical Framework

Sustainability refers to “... the endurance of systems and processes”. In particular debt sustainability is taken to mean the avoidance of default and is achieved where the present value of the primary balance is equal to the current debt – where the government’s inter-temporal budget constraint (IBC) is satisfied. We adopt a simple theoretical framework advanced by Bohn (1998) to model the fiscal response of the primary balance to changes in the debt-to-GDP ratio.

The government’s IBC takes the standard form:

$$D_{t+n} = (1 + r)^n \cdot D_t - \sum_{j=0}^n (1 + r)^{n-j} \cdot S_{t+j} \quad (1)$$

Where (D) is the stock of debt and (S) is the primary balance with subscripts denoting the corresponding time period for the respective variables. Equation (1) is a simple accounting equation that explains the evolution of government debt – as a function of the interest cost and an error correction term captured by the primary surplus. For a given interest rate, governments can reduce their debt by running sustained primary surpluses.

Solving for the current stock of debt and taking expectations of future values yield:

$$D_t = (1 + r)^{-n} \cdot E_t (D_{t+n}) + \sum_{j=0}^n (1 + r)^{-j} \cdot E_t (S_{t+j}) \quad (2)$$

⁴ Lukkezen and Rojas-Romagosa (2012) show that greater uncertainty in developing countries result in lower tolerance on the part of creditors which in turn forces these countries to respond much more strongly in an effort to stabilize debt.

We follow the standard approach in the literature and express equation (2) as a share of GDP in equation (3). This normalizes debt across countries and allows for cross country analyses.

$$d_t = \left(\frac{1+g}{1+r}\right)^n \cdot E_t(d_{t+n}) + \sum_{j=0}^n \left(\frac{1+g}{1+r}\right)^j \cdot E_t(s_{t+j}) \quad (3)$$

Taking the limit as $n \rightarrow \infty$ leaves us with two key conditions⁵:

$$\lim_{n \rightarrow \infty} \left(\frac{1+g}{1+r}\right)^n \cdot E_t(d_{t+n}) \leq 0 \quad (4)$$

$$d_t = \sum_{j=0}^{\infty} \left(\frac{1+g}{1+r}\right)^j \cdot E_t(s_{t+j}) \quad (5)$$

Equations (4) and (5) give the transversality (no ponzi) condition and the government's inter-temporal budget constraint, respectively. The former assumes that governments cannot hold or accumulate debt in perpetuity. A violation of the transversality condition would suggest that the government is unable to repay its initial debt, a situation that would cause rational agents to withhold credit and therefore serves as a real constraint on government's debt accumulation behavior.⁶

Comparative statics from equation (5) captures several important relationships. Firstly, higher rates of growth in real economic output, sustained over time, can support higher levels of debt. Secondly, access to low cost financing can significantly improve governments' ability to sustain higher levels of debt.⁷ Thirdly, higher levels of debt can be sustained if governments can

⁵ Equations (4) and (5) assume that GDP grows at a constant rate g and thus $y_{t+n} = (1+g)^n y_t$.

⁶ Bohn (2008) argues that there are plausible conditions under which (4) may not have defined limits or may have a negative limit, but that such considerations will unnecessarily serve to complicate the analysis. We follow the standard approach in the literature and assume that equation (4) holds at equality.

⁷ Whilst low interest costs on sovereign debt can come from simply borrowing at more favorable interest rates, another avenue is through seigniorage or inflation. See Reinhart and Sbrancia (2011) for a discussion on the use of financial repression as a way to avoid having to make real fiscal policy adjustments.

consistently run positive primary balances. Notably, if high debt is associated with higher financing costs and lower economic growth rates, the theory implies a need for even larger fiscal responses to ensure stabilization. This has prompted a number of studies to suggest that a simple positive linear relationship between primary balance and debt-to-GDP might not be sufficient (Ghosh, Kim, Mendoza, Ostry, & Qureshi, 2011) to ensure sustainability. These studies point to a need for stricter sustainability criteria, namely that the systematic policy response – the coefficient on the primary balance – is greater than the interest rate - growth rate differential ($r - g$).

Whether countries are able to respond in the manner suggested is an open empirical question - one we examine in this essay. We might expect fiscally responsible governments to make larger adjustments when debt levels are high. However, at higher levels of debt, larger increases in the primary balance may be economically, socially and politically difficult – a situation Gosh et al. (2011) labels ‘fiscal fatigue’. The credibility of the overall fiscal adjustment plan therefore hinges on the practicality of the path for the primary balance and governments’ ability to always react to increases in debt. How countries react is also important – do they increase revenues or cut primary spending. The margin of response adopted can have broader economic and welfare implications. For example Krogstrup (2002) points out that tax increases in a context of high tax competition can introduce potentially large distortions that compromise economic welfare. Additionally, the margin of response may also vary depending on the level of debt. For example, if countries are averse to tax increases they might choose spending cuts when debt is relatively low and rely on tax hikes as a last resort when debt is relatively high.

Literature Review

Much of the recent discussion in the literature begins with Bohn (1998) who examined the fiscal response to debt for the US using time series methods. He finds that after controlling for volatility in output and government spending (due largely to wartime spending), the primary balance in the US bore a consistently positive relationship with the debt to-GDP-ratio. However, Bohn's work was focused on the US and so the question of how other governments react to debt remained largely unexplored. Subsequently, several studies have applied Bohn's framework to examine the FRF in other countries. Lukkezen and Rojas-Romosagosa (2012) use time series analysis to examine the fiscal response to debt for seven OECD countries. They find that the US, UK, Netherlands and Belgium responded to growing debt by increasing their primary balance, but Spain, Portugal and Iceland did not. They find nonlinearities in the relationship between the primary balance and debt-to-GDP going in both directions. The UK and the Netherlands respond more strongly – increase their primary balance by more at higher levels of debt. However, Spain, Portugal and Iceland struggle make the necessary adjustments, and primary balance actually declines at relatively high levels of debt.

Other studies approach this question using panel data methods. Celasun and Ostry (2006) examined the fiscal response for 34 emerging market economies over the period 1990 - 2004 using limited information maximum likelihood (LILM) and system GMM models. They find a positive relationship between primary balance and debt-to-GDP. Gosh et al (2011) examine a similar question for a panel of advanced economies over the period 1970 – 2007 and also find generally positive responses in the primary balance. They also report nonlinearities in the relationship between debt and the primary balance. In particular, they find that the response of the primary balance start to decline at debt-to-GDP of about 90 –100 percent and turns negative

as it approaches 150 percent. This finding is counter to Bohn (1998, 2008) who found stronger responses at higher levels of debt for the US. Mendoza and Ostry (2008) examine the fiscal reaction function for a group of 22 industrial countries (IC) and 33 emerging economies (EE).⁸ Interestingly this paper offers a direct comparison of the fiscal response for two distinct country groupings. They find that IC have higher debt ratios than EE – in line with the debt intolerance conjecture (Reinhart, Rogoff & Savastano, 2003). They employ a fixed effects model with errors corrected for first order autocorrelation and find a positive relationship between the primary balance and debt-to GDP for both IC and EE. They also report that the response among EE was much stronger than in IC – a 1 percentage point increase in the debt ratio was associated with .036 and 0.02 of a percentage point increase in the primary balance for EE and IC respectively. Tests for nonlinearities did not yield significant results for IC but in the case of EE showed a weakening in response when debt-to-GDP exceeded 50 percent.⁹

Afonso and Hauptmeier (2009) expand the analysis to account for effects of fiscal rules and level of government decentralization in estimating the response of the primary balance to debt-to-GDP ratio for a panel of 27 EU member countries. In general, their results corroborate earlier findings of a positive response to debt. They report smaller response magnitudes when the debt ratio is greater than 80 percent. Interestingly, they also examine the response of the primary spending component of the primary balance and find no significant effect. This result suggests that the EU countries studied respond largely through the revenue margin.¹⁰

⁸ Estimates for industrial countries were done using data from 1970 – 2005 and in the case of developing countries from 1990-2005.

⁹ One test they did was to split the sample into high and low debt countries and run separate regressions and compared the estimated coefficients. They found that for highly indebted EM countries the coefficient on the debt-to-GDP was not statistically significant.

¹⁰ They find that the existence of fiscal rules and public spending decentralization has a positive effect on the primary balance.

In work closely related to this essay, Krogstrup (2002) examines the response of taxes and spending to debt-to-GDP for a panel of EU countries. She finds that countries respond along both the tax and spending margins, with slightly larger adjustments for the latter. She also finds that debt servicing in the context of high capital mobility does not affect countries response along the tax margin but that countries respond by significantly reducing spending. These results highlight some interesting patterns in fiscal policy response for the EU. Firstly, countries respond along both the tax and spending margins and secondly, there is a tendency for larger responses along the spending margin, particularly when there is tax competition. Reicher (2013) estimates the reaction of different categories of government spending, taxes and transfers to debt-to-GDP for a panel of 20 industrial countries. In contrast to Krogstrup (2002), he finds that in general, the countries studied increase their primary balance in response to debt and that much of this increase comes through higher taxation with much smaller adjustments in government spending.¹¹

In sum, the existing literature places inordinate focus on examining the fiscal response among industrial countries and countries within the EU. Very few studies go beyond the conventional approach and actually disaggregate the fiscal response to better understand the anatomy of debt consolidation within these countries. This essay attempts to fill this gap by examining the response of the primary balance and its structural components, for a sample of developing countries.

¹¹ Estimates for individual tax measures studied (labor tax rate, capital tax rate and tax as a share of GDP) were all positive but in general were imprecisely estimated.

Data

We examine the fiscal response to debt-to-GDP for a panel of 54 developing countries.¹² In an attempt to treat with concerns about the short time span over which fiscal policy analyses in less developed countries have generally been conducted, we use data over 22 years, from 1990 to 2011.

The fiscal variables used in this research are from the International Monetary Fund (IMF) Government Finance Statistics (GFS) and we use data for the consolidated central government. The use of consolidated data reduces the likelihood of double counting of fiscal aggregates across the various levels of government within countries. It also mitigates potential comparability problems across countries with different governmental structures. The choice of the level of governmental aggregation is informed largely by data availability. In developing countries, fiscal reporting at the subnational level is generally weak, and so, much of the available data are for central government operations. Additionally, because major fiscal policy changes are enacted at the central government level in most developing countries we feel that the choice of the level aggregation is further justified.¹³ Table 1 presents summary statistics for the variables used.

¹² See Table A.1 in Appendix A for a list of the countries used in this research.

¹³ Devarajan, Swaroop and Zou (1996) examine the effect of government spending on growth using a subset of countries with both general and central government fiscal data and found similar results. This suggests that the level of aggregation may not significantly affect our results.

Table 1: Summary Statistics, 1990 – 2011

Variables	No Obs.	Mean	Std. Dev.	Min	Max
Primary Balance	940	0.033	0.032	-0.131	0.181
Fiscal Effort	953	-0.036	0.056	-0.271	0.083
Central Govt. Revenue	966	0.229	0.077	0.077	0.571
Central Govt. Tax Revenue	978	0.160	0.050	0.048	0.318
Central Govt. Primary Spending	953	0.196	0.077	0.050	0.483
Lag General Govt. Debt / GDP	1105	0.531	0.301	0.010	2.896
Output Gap	1183	-0.003	0.053	-0.462	0.253
Expenditure Gap	972	-0.004	0.062	-0.401	0.315
Income Tax / Total Tax Revenue	992	0.307	0.136	0.000	0.850
Infation Rate	1174	0.629	5.548	-0.060	154.444
Openness	1164	0.771	0.357	0.108	2.204
Net ODI Share in GDP	1183	0.028	0.042	-0.007	0.261
Government Effectiveness	1042	5.273	4.950	-8.000	10.000
IMF or Fiscal Rule	1166	0.591	0.492	0.000	1.000
Agriculture Share in GDP	1110	0.151	0.107	0.023	0.660

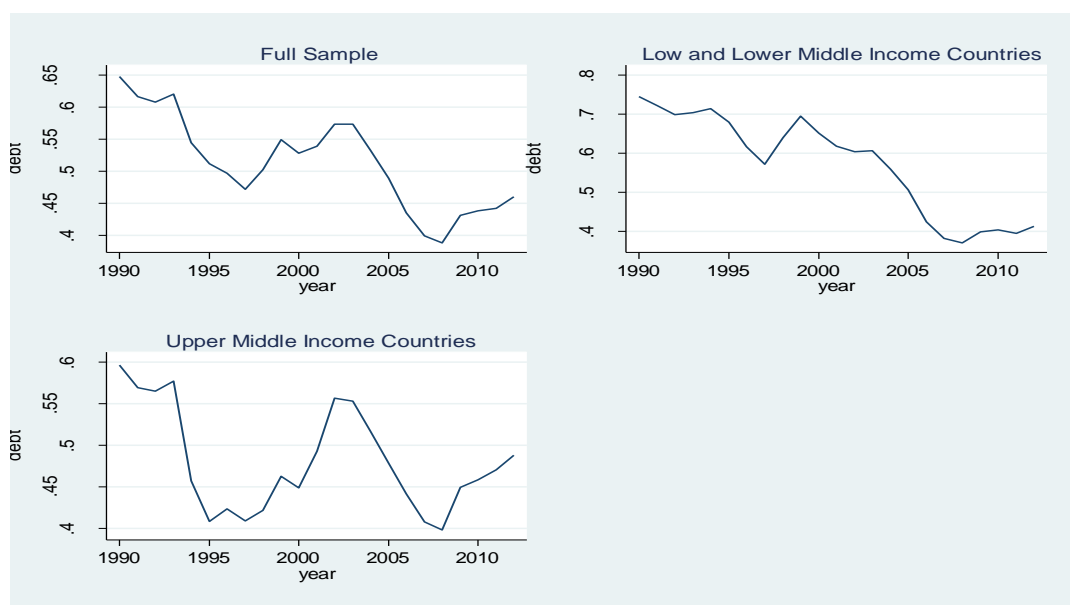
For the dependent variables, we examine five measures of fiscal response – primary balance, ‘fiscal effort, general revenues, tax revenues and primary spending. Essentially this amounts to decomposing the primary balance and ‘fiscal effort’ and examining the response of each component. All outcomes examined are expressed as a share of GDP.

The primary variable of interest is the lagged debt-to-GDP ratio and is taken from the IMF historical public debt data base.¹⁴ Figure 1 shows the trend in general government average debt ratios for the countries in the full sample as well as low / lower-middle income countries and high / upper-middle income countries, over the period 1990 – 2011. Though the average debt ratio displays some variability, countries reduced debt-to-GDP by about 18 percentage points over the period. The downward trend in debt-to-GDP appears to be driven largely by low /

¹⁴ Data on debt-to-GDP were compiled by S. M. Ali Abbas, Nazim Belhocine, Asmaa El-Ganainy and Mark Horton and measures general government debt-to-GDP. The general government sector includes all government units and all non-market nonprofit institutions that are controlled and mainly financed by government units, containing central and local governments. It does not include public corporations or quasi-corporations. Whilst the data series aimed to capture general government debt, the lack of public debt data at this level in many countries meant that central government debt data were used in these instances. For more details see Abbas et al (2010).

lower-middle income countries, where the reduction was about 33 percentage points compared to 11 percentage points for high / upper-middle income countries.

Figure 1: Debt-to-GDP Ratio, 1990 – 2011



Notes: Figure shows debt dynamics by country grouping over the period 1990-2011. Compiled using data on general government debt-to-GDP from S. M. Ali Abbas, Nazim Belhocine, Asmaa El-Ganainy and Mark Horton.

Figures A.1 – A.3 in Appendix A present scatter plots of the various fiscal outcomes against debt-to-GDP for the full sample, low / lower-middle and high / upper-middle income countries, respectively. The plots show a generally positive correlation between the primary balance and fiscal effort with lagged debt-to-GDP for all country groupings. The plots also highlight relatively stronger responses for general revenues and tax revenues for upper-middle income compared to low / lower-middle income countries; and low / lower-middle income countries show a slightly negative relationship between primary spending and debt. These relationships are subject to more formal empirical examination below.

The set of covariates used include a measure for the output gap, expenditure gap, inflation, trade openness, share of agriculture in GDP, share of income tax in total tax revenue, a measure of institutional democracy (Polity IV), net overseas development assistance (ODA) as a share of GNI¹⁵, an indicator for whether the country had a fiscal rule and one for if these us an existing IMF agreement. A description of all the variables used along with their sources is presented in Table A.2 of Appendix A.

Empirical Strategy

The estimation strategy adopts a dynamic panel regression approach. We first examine a baseline FRF of the form:

$$y_{i,t} = \beta_0 + \beta_1 y_{i,t-1} + \beta_2 g_t + \beta_3 e_t + \beta_4 debt_{i,t-1} + x'_{it} \beta_5 + \tau_i + \gamma_t + \varepsilon_{i,t} \quad (6)$$

where i indexes the countries in the panel and t the time period over which the analyses are conducted.

The dependent variable, $y_{i,t}$ captures the government's fiscal response and is measured along the various margins outlined above. The contemporaneous fiscal response may depend on the previous year's fiscal balance(s) and thus implies the need for a dynamic model.

Accordingly, all specifications of the model include a lagged dependent variable $y_{i,t-1}$, that captures the dynamic effects of the fiscal adjustment process within countries. The independent variable of interest is the lagged debt-to-GDP ratio ($debt_{i,t-1}$). Following the empirical literature e_t and g_t are measures of the government expenditure gap and the real output gap respectively.

The other controls used are included in x_{it} .

¹⁵ This measure captures official aid flows and do not include for example debt forgiveness that is often extended to highly indebted poor countries (HIPC). It therefore serves as lower bound of development financing / budgetary support that some countries may be able to access.

We also test for possible nonlinearities in the relationship between lag debt-to-GDP and each fiscal response outcome. We do this first by fitting a quadratic specification of the baseline model. Additionally, we consider an alternative specification of the baseline model in equation (6) and estimate linear spline regressions. This is useful since the nature on the nonlinearity may not follow a U shape or inverse U shape, but could instead demonstrate a kinked relationship.

$$y_{i,t} = \beta_0 + \beta_1 y_{i,t-1} + \beta_2 \text{debt}_{i,t-1} * I(\text{debt}_{i,t-1} \leq D) + \beta_3 \text{debt}_{i,t-1} * I(\text{debt}_{i,t-1} > D) + \beta_4 x_{it} + \tau_i + \gamma_t + \varepsilon_{i,t} \quad (7)$$

Where D is some debt threshold and $I(\cdot)$ is an indicator variable that takes the value of one if the lagged value of debt is greater than the chosen threshold and zero otherwise. Equation (7) allows for one effect for debt ratios less than the specified threshold (D) and another effect when debt is greater than the D.¹⁶

Estimation of equations (6) and (7) must take into account possible endogeneity of key independent variables. In particular, lagged debt-to-GDP may be correlated with unobserved time invariant country characteristics captured in $\varepsilon_{i,t}$. As highlighted by Celasun, Debrun and Ostry (2006) a country's debt is an accumulation of previous deficits, if there are unobserved factors that causes it to generate large primary surpluses - relatively low levels of debt - this will downward bias the estimated coefficient on the debt ratio. A second potential source of endogeneity comes from the persistence of policy shocks. For example, a fiscal policy shock in the previous period that is persistent – implying that $\varepsilon_{i,t-1}$ and $\varepsilon_{i,t}$ are serially correlated, will render the estimates inconsistent since we know that $\text{debt}_{i,t-1}$ is correlated with that initial shock.

¹⁶ We re-estimated equation (7) and include an intercept dummy and got similar results to the baseline specification.

To mitigate concerns regarding the possible endogeneity of debt-to-GDP and other key covariates we first estimate a fixed effects (FE) panel regression model to control for possible time invariant country specific heterogeneity.¹⁷ However the FE model does not deal with the potential endogeneity that comes through the serial correlation of the errors. To mitigate this potential source of endogeneity we adopt a dynamic difference generalized methods of moments (DGMM) model with forward orthogonal deviation (FOD) (Arellano & Bover, 1995).¹⁸ The DGMM has several features that make it particularly useful in our context. Firstly, it works well for panels with a short time dimension - small (T) and a relatively large number of panel units (N).¹⁹ Secondly, DGMM with FOD transforms the data by subtracting the average of all available future observations from their contemporaneous values to purge the fixed effects. Relative to first difference transformations, this approach minimizes data loss for unbalanced panels. Thirdly it uses instruments for the potentially endogenous variables that are drawn from within the data and thereby circumvents the difficulty of having to find valid external instruments. In particular, the DGMM uses the lagged transformation of the potentially endogenous variables as instruments.

Results

Results for the Full Sample

Table 2 summarizes the results for the primary balance, from FE and DGMM regression, with and without year dummies. The preferred model is the DGMM with year dummies.²⁰ FE

¹⁷ The Hausman test for random versus fixed effects rejects the null, in support of a fixed effects specification.

¹⁸ FOD minimizes data loss in unbalanced panels; and since lagged observations are not used in the transformation they are valid instruments.

¹⁹ Celasun and Kang (2006) advance that if T is relatively large (7 – 8) then the number of instruments generated will be huge if unrestricted and can weaken the Sargan and Hansen J tests of instrument validity.

²⁰ The inclusion of year dummies in the DGMM model can lead to instrument proliferation which can significantly weaken the Hansen-J test of over-identifying restrictions. Notwithstanding, we take comfort that the Hansen-J p-values from DGMM regressions without year dummies suggest that collectively the instruments are valid.

estimates (columns 1-2) suggest that a one percentage point increase in lagged debt-to-GDP is associated with a 0.026 and 0.022 percentage point increase in the primary balance, for regressions with and without time dummies, respectively. DGMM estimates (columns 3-4) are a little larger, with coefficient estimates of 0.043 and 0.034, with and without time dummies, respectively.

Table 2: Results for Primary Balance

	Primary Balance			
	FE (1)	FE (2)	DGMM (3)	DGMM (4)
Lag Dependent Variable	0.499*** (0.0378)	0.466*** (0.0418)	0.588*** (0.0786)	0.559*** (0.0723)
Lag Debt/GDP	0.0262*** (0.00600)	0.0219*** (0.00623)	0.0429*** (0.0119)	0.0337*** (0.0102)
Output Gap	0.0883*** (0.0229)	0.0577** (0.0254)	0.0893*** (0.0226)	0.0598** (0.0233)
Expenditure Gap	-0.0333* (0.0170)	-0.0297* (0.0157)	-0.0197 (0.0182)	-0.0206 (0.0190)
Inflation	0.00112*** (0.000216)	0.000784** (0.000338)	0.000872*** (0.000224)	0.000656*** (0.000303)
Trade Openness	0.0126 (0.0111)	0.00596 (0.00926)	0.0121 (0.0111)	0.00446 (0.00881)
Share of Agriculture in GDP	0.0767** (0.0359)	0.0403 (0.0469)	0.0330 (0.0410)	0.0241 (0.0455)
Share - Income Tax	-0.0293 (0.0180)	-0.0145 (0.0159)	-0.0260 (0.0183)	-0.0142 (0.0154)
Polity IV	-0.00120*** (0.000339)	-0.00106*** (0.000338)	-0.00102*** (0.000345)	-0.000951*** (0.000317)
Constant	-0.00272 (0.0114)	0.00999 (0.0118)		
Observations	632	632	588	588
R-squared	0.437	0.494		
No. Countries	44	44	43	43
AR(2) (p-value)			0.129	0.261
Hansen J (p-value)			0.340	0.994
Year Dummy	no	yes	no	yes

Notes: Dependent variable is the primary balance expressed as a share of GDP. Results are for Fixed Effects (FE) and difference GMM (DGMM) regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995). To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables. Arellano-Bond AR (2) diagnostic tests suggest no second order serial correlation in the errors. Hansen J (p-value) suggests that our instruments are valid. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Additionally the coefficient estimates in both DGMM regressions with and without year dummies are qualitatively similar for all outcomes examined.

We find similar results for fiscal effort, but with slightly smaller coefficient estimates, presented in Table 3. FE estimates (columns 1-2) suggest that a one percentage point increase in lagged debt-to-GDP is associated with a 0.019 and 0.015 percentage point increase in the fiscal effort, for regressions with and without time dummies respectively. Like the primary balance, DGMM estimates (columns 3-4) are larger, with coefficient estimates of 0.036 and 0.023, with and without time dummies, respectively.

Table 3: Results for the Fiscal Effort

	Fiscal Effort			
	FE (1)	FE (2)	DGMM (3)	DGMM (4)
Lag Dependent Variable	0.497*** (0.0937)	0.482*** (0.0981)	0.547*** (0.133)	0.545*** (0.148)
Lag Debt/GDP	0.0193*** (0.00622)	0.0152** (0.00681)	0.0360*** (0.0126)	0.0226** (0.0109)
Output Gap	0.103*** (0.0265)	0.0618* (0.0323)	0.107*** (0.0249)	0.0610** (0.0296)
Expenditure Gap	-0.0329** (0.0144)	-0.0320** (0.0138)	-0.0190 (0.0144)	-0.0258 (0.0160)
Inflation	-0.000705** (0.000291)	-0.00112*** (0.000332)	-0.000725** (0.000301)	-0.00104*** (0.000390)
Trade Openness	0.0106 (0.0146)	0.00494 (0.0120)	0.0110 (0.0149)	0.00493 (0.0113)
Share of Agriculture in GDP	0.177*** (0.0604)	0.0951 (0.0575)	0.133* (0.0715)	0.0774 (0.0564)
Share - Income Tax	0.0314 (0.0304)	0.0519 (0.0314)	0.0321 (0.0301)	0.0490 (0.0309)
Polity IV	-0.000794** (0.000366)	-0.000624* (0.000310)	-0.000697* (0.000400)	-0.000574* (0.000299)
Constant	-0.0678*** (0.0207)	-0.0567*** (0.0202)		
Observations	632	632	588	588
R-squared	0.427	0.480		
No. Countries	44	44	43	43
AR(2) (p-value)			0.165	0.234
Hansen J (p-value)			0.302	0.999
Year Dummy	no	yes	no	yes

Notes: Dependent variable is fiscal effort - defined as tax revenues minus primary spending, and is expressed as a share of GDP. Results are for Fixed Effects (FE) and difference GMM (DGMM) regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995). To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables. Arellano-Bond AR (2) diagnostic tests suggest no second order serial correlation in the errors. Hansen J (p-value) suggests that our instruments are valid. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

In general, the results are indicative of responsible fiscal policy for the sampled countries. The response of the primary balance and the fiscal effort, however, do not tell us the margins along which fiscal policy adjustments are made – whether countries tend to rely more heavily on revenue increases or spending cuts, or both. Table 4 summarizes results for the fiscal components from DGMM regressions. Both revenue outcomes have a positive and statistically significant relationship with lag debt-to-GDP. A one percentage point increase in lag debt-to-GDP is associated with an increase of between 0.022 – 0.027 (columns 1-2) percentage points in total revenues, for regressions with and without year dummies, respectively. This compares to the results for tax revenues, with slightly smaller coefficient estimates of 0.017 and 0.014 (columns 3-4) percentage points. Results for primary spending indicate a negative and statistically significant relationship in the order of -0.022 and -0.013 percentage points with and without year dummies respectively.

The results suggest that countries adjust along both the revenue and expenditure margins. Baldacci et al. (2014) rationalizes these findings, arguing that fiscal adjustments along both tax and spending margins when debt is high and credit constraints are about binding can be more favorable for economic growth post consolidation. This is plausible if individuals have diminishing marginal utility in net private income and public spending, then spreading the adjustments across the tax and expenditure margin can minimize welfare loss.

Table 4: Results for Fiscal Components

	Total Revenue		Tax Revenue		Primary Spending	
	DGMM (1)	DGMM (2)	DGMM (3)	DGMM (4)	DGMM (5)	DGMM (6)
Lag Dependent Variable	0.609*** (0.126)	0.479*** (0.163)	0.542*** (0.111)	0.554*** (0.0967)	0.597*** (0.102)	0.685*** (0.0753)
Lag Debt/GDP	0.0220** (0.0104)	0.0272** (0.0131)	0.0174** (0.00764)	0.0143** (0.00602)	-0.0222** (0.0107)	-0.0127* (0.00698)
Output Gap	0.0452* (0.0264)	0.0296 (0.0343)	0.0558*** (0.0200)	0.0232 (0.0216)	-0.0477* (0.0256)	-0.0215 (0.0274)
Expenditure Gap	-0.00194 (0.0188)	0.0143 (0.0191)	0.00826 (0.0178)	0.0139 (0.0181)	0.0206 (0.0168)	0.0291* (0.0169)
Inflation	0.000549 (0.000563)	0.000568 (0.000642)	-0.00108*** (0.000391)	-0.00117** (0.000485)	-0.000233 (0.000483)	-1.79e-06 (0.000457)
Trade Openness	0.0230** (0.0111)	0.0190 (0.0121)	0.0235*** (0.00875)	0.0155** (0.00761)	0.0107 (0.0137)	0.00830 (0.0102)
Share of Agriculture in GDP	-0.100** (0.0466)	-0.0654 (0.0632)	-0.0134 (0.0314)	0.00293 (0.0337)	-0.141*** (0.0494)	-0.0784* (0.0407)
Share - Income Tax	-0.0360 (0.0318)	-0.0415 (0.0337)	0.0250 (0.0191)	0.0217 (0.0164)	-0.0101 (0.0284)	-0.0269 (0.0272)
Polity IV	-0.000501 (0.000352)	-0.000589 (0.000473)	-0.000114 (0.000338)	-0.000136 (0.000324)	0.000547 (0.000377)	0.000407 (0.000333)
Observations	612	612	611	611	588	588
No. Countries	43	43	43	43	43	43
AR(2) (p-value)	0.740	0.518	0.540	0.664	0.977	0.892
Hansen J (p-value)	0.486	0.980	0.443	0.998	0.328	1
Year Dummy	no	yes	no	yes	no	yes

Notes: Dependent variables are general revenue, tax revenue and primary spending, all expressed a share of GDP. Results are for DGMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995). To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables. Arellano-Bond AR (2) diagnostic tests suggest no second order serial correlation in the errors. Hansen J (p-value) suggests that our instruments are valid. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

For the controls, we discuss results for the preferred model only. The estimated coefficient for the output gap is positive and statistically significant in the primary balance and fiscal effort regressions only. Additionally, it bears a positive correlation with total revenues and tax revenues and is negatively correlated with primary spending - though not statistically significant for any. This provides suggestive evidence of countercyclical fiscal policy in the sampled countries, which is counter to previous literature that finds generally pro-cyclical fiscal reactions in developing countries.²¹ The expenditure gap is found to be negatively correlated with all outcomes except primary spending. As expected higher spending relative to trend is

²¹ Celasun et al. (2006) point out that this result should be interpreted with caution as it may be driven by deteriorating fiscal outturns which are natural consequence of economic slowdowns and recessions.

positively and statistically significantly correlated with the primary spending. Inflation is positively and significantly correlated with the primary balance, but negatively correlated with fiscal effort and tax revenues. However the point estimate is small and suggests a relatively insignificant economic relationship. Trade openness is not found to significantly affect the primary balance, fiscal effort or primary spending, but is positively correlated with revenues, and particularly with tax revenues. The share of agriculture in GDP is negatively and statistically significantly correlated with primary spending. The coefficients on our measure of institutional democracy (Polity IV) in the primary balance and fiscal effort regressions are negative - suggesting that more democratic countries are less fiscally responsible than more democratic countries.

Non Linear Fiscal Reaction

We also test for potential nonlinearities in the fiscal response and report our results in Table 5. We do this first by fitting a quadratic regression of the baseline model and secondly by estimating separate spline regressions with knots at debt-to-GDP ratios of 50 and 90 percent.²² Results for the primary balance are presented in columns 1-3. For the quadratic model (column 1) the coefficient on lag debt-to-GDP though negative is statistically insignificant and the coefficient on the quadratic term is positive and statistically significant. This suggests that countries in the sample respond more strongly through increases in the primary balance at higher levels of debt. Results from spline regressions (columns 2 and 3) confirm these findings with insignificant effects at debt ratios below the specified knots and positive and statistically significant coefficients for debt ratios above the knots. However, this evidence is suggestive at

²² The knots were chosen based on findings from previous literature that identified nonlinearities in the relationship the primary balance at debt ratios of 50 and 90 percent of GDP (Celasun et al., 2006; Mondoza and Ostry, 2007; Gosh et al., 2011).

best, since Wald tests for equality of slope coefficients above and below the knots fail to reject the null of equality of coefficients. We find qualitatively similar results for the fiscal effort (columns 4-6) – negative correlation at relatively low levels of debt and positive correlations at relatively high levels of debt. However, the lack of statistical significance for the key variables of interest is not suggestive of a nonlinear relationship.

Results for the fiscal components are particularly interesting. For total revenues (columns 7 - 9) we find null effects in the quadratic model. However, for the spline regressions, we find a negative but statistically insignificant response (-0.089) when debt-to-GDP is below 50 percent, and a positive and statistically significant response (0.045) above 50 percent. Wald tests reject the null hypothesis of equal slopes on either side of the knot – suggesting that countries increase total revenues more aggressively when debt is above 50 percent of GDP compared to when debt is below 50 percent of GDP. Point estimates display a similar pattern at the 90 percent knot, except here the Wald test fails to detect statistically significant differences in the estimated coefficients on either side of the ‘knot’.

The nonlinearities are more pronounced for tax revenues (column 10). Results from the quadratic model indicate a negative and statistically significant relationship with lag debt-to-GDP but a positive and highly significant relationship with the quadratic term. Results from the 50 percent spline regression indicate a negative though statistically insignificant response (-0.046) when debt is below 50 percent of GDP, but a positive and statistically significant response (0.024) when debt is above 50 percent of GDP. In regressions that use a debt knot at 90 percent of GDP, we find similar results - a negative though statistically insignificant response (-0.010) when debt is below 90 percent of GDP, but a positive and statistically significant response (0.027) when debt is above 90 percent of GDP. Wald tests confirm that the slope coefficients on

either side of the knots are statically different. For primary spending we find no significant effect in either the quadratic or 90 percent spline specification, but find evidence of a kinked relationship at debt-to-GDP of 50 percent. In particular, the results in (column 14) indicate a negative and statistically significant response (-0.123) when the debt ratio is below 50 percent and a positive though statistically insignificant response when debt is above 50 percent of GDP.

Table 5: Results for Non Linear Fiscal Response

	Primary Balance			Fiscal Effort			Total Revenue			Tax Revenue			Primary Expenditure		
	Quadratic (1)	Spline 50% (2)	Spline 90% (3)	Quadratic (4)	Spline 50% (5)	Spline 90% (6)	Quadratic (7)	Spline 50% (8)	Spline 90% (9)	Quadratic (10)	Spline 50% (11)	Spline 90% (12)	Quadratic (13)	Spline 50% (14)	Spline 90% (15)
Lag Dependent Variable	0.589*** (0.0801)	0.536*** (0.0735)	0.543*** (0.0745)	0.608*** (0.168)	0.571*** (0.144)	0.579*** (0.154)	0.413** (0.190)	0.429*** (0.153)	0.440*** (0.159)	0.529*** (0.0941)	0.468*** (0.102)	0.494*** (0.0976)	0.731*** (0.0890)	0.661*** (0.0810)	0.669*** (0.0767)
Lag Debt/GDP	-0.0134 (0.0246)			-0.0428 (0.0435)			-0.00240 (0.0369)			-0.0281* (0.0157)			-0.0381 (0.0269)		
Lag Debt/GDP Squared	0.0188* (0.0102)			0.0166 (0.0144)			0.0154 (0.0134)			0.0185*** (0.00508)			0.00883 (0.00927)		
Debt/GDP below 50%		-0.00987 (0.0544)			-0.0148 (0.0685)			-0.0892 (0.0679)			-0.0462 (0.0376)			-0.123** (0.0516)	
Debt/GDP above 50%		0.0409*** (0.0128)			0.000745 (0.0142)			0.0454*** (0.0175)			0.0236*** (0.00811)			0.00209 (0.0102)	
Debt/GDP below 90%			0.00177 (0.0205)			-0.0386 (0.0387)			0.00242 (0.0236)			-0.00969 (0.0138)			-0.0238 (0.0196)
Debt/GDP above 90%			0.0503*** (0.0159)			0.0171 (0.0156)			0.0401* (0.0211)			0.0271*** (0.00844)			-0.00915 (0.0114)
Output Gap	0.0520** (0.0245)	0.0571** (0.0253)	0.0544** (0.0230)	0.0503 (0.0496)	0.0576 (0.0446)	0.0481 (0.0452)	0.0204 (0.0345)	0.00924 (0.0395)	0.0218 (0.0355)	0.0181 (0.0210)	0.0133 (0.0240)	0.0162 (0.0226)	-0.0122 (0.0317)	-0.0454 (0.0327)	-0.0285 (0.0282)
Expenditure Gap	-0.0295* (0.0178)	-0.0208 (0.0178)	-0.0261 (0.0167)	-0.0603*** (0.0164)	-0.0510*** (0.0168)	-0.0558*** (0.0148)	0.0153 (0.0170)	0.0226 (0.0173)	0.0142 (0.0168)	0.0108 (0.0146)	0.0191 (0.0170)	0.0143 (0.0167)	0.0230 (0.0176)	0.0338** (0.0167)	0.0291* (0.0167)
Inflation	0.000770** (0.000326)	0.000704** (0.000276)	0.000875*** (0.000278)	-0.00250*** (0.000615)	-0.00253*** (0.000385)	-0.00232*** (0.000472)	0.000543 (0.000823)	0.000655 (0.000582)	0.000757 (0.000616)	-0.00110** (0.000535)	-0.00129*** (0.000433)	-0.00112** (0.000451)	0.000293 (0.000350)	2.78e-05 (0.000426)	6.06e-05 (0.000476)
Trade Openness	0.00550 (0.00874)	0.00684 (0.00917)	0.00534 (0.00812)	-0.000668 (0.0142)	-0.00127 (0.0146)	-0.00171 (0.0124)	0.0195 (0.0119)	0.0243* (0.0144)	0.0199 (0.0123)	0.0161** (0.00687)	0.0182** (0.00740)	0.0158** (0.00716)	0.00965 (0.0121)	0.0116 (0.0167)	0.00765 (0.0110)
Agriculture/GDP	0.0520 (0.0480)	0.0448 (0.0517)	0.0508 (0.0462)	0.187** (0.0947)	0.172** (0.0848)	0.196** (0.0872)	-0.0631 (0.0692)	-0.00322 (0.0701)	-0.0376 (0.0573)	0.0340 (0.0349)	0.0436 (0.0377)	0.0338 (0.0336)	-0.0434 (0.0451)	-0.0274 (0.0557)	-0.0694 (0.0439)
Share - Income Tax	-0.0183 (0.0547)	-0.0221 (0.0181)	-0.0229 (0.0153)	0.0844 (0.0995)	0.0738*** (0.0243)	0.0666** (0.0259)	0.0120 (0.0591)	-0.0620* (0.0371)	-0.0489 (0.0346)	0.0285 (0.0500)	0.0154 (0.0183)	0.0179 (0.0173)	-0.0831* (0.0502)	-0.0448 (0.0330)	-0.0290 (0.0286)
Polity IV	-0.000984*** (0.000295)	-0.000975*** (0.000264)	-0.00101*** (0.000291)	-0.000376 (0.000346)	-0.000358 (0.000326)	-0.000400 (0.000293)	-0.000576 (0.000560)	-0.000610 (0.000458)	-0.000601 (0.000484)	-0.000177 (0.000353)	-0.000142 (0.000326)	-0.000140 (0.000336)	0.000275 (0.000332)	0.000443 (0.000414)	0.000441 (0.000339)
Observations	588	590	590	588	590	590	612	614	614	611	613	613	588	590	590
No of Countries	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43
AR 2 (p-value)	0.415	0.236	0.258	0.136	0.099	0.095	0.239	0.432	0.388	0.550	0.836	0.865	0.839	0.846	0.934
Hnasen J (p-value)	0.981	0.990	0.981	0.995	0.999	0.998	0.979	0.999	0.971	0.995	0.995	0.996	1	1	1
Wald Test (p-value)		0.411	0.115		0.837	0.261		0.078	0.298		0.099	0.045		0.021	0.582
Year Dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: Dependent variables are primary balance, ‘fiscal effort’, general revenue, tax revenue and primary spending, all expressed a share of GDP. Results are for DGMM regressions using forward orthogonal deviations (FOD) proposed by Arellano and Bover (1995). To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables. Hansen J (p-values) are implausibly large suggesting weakened diagnostics from instrument proliferation. In regression that exclude year dummies the Hansen J (p-values) are in normal ranges and suggest that our instruments are collectively valid. Arellano-Bond AR (2) diagnostic tests suggest no second order serial correlation in the errors. Robust standard errors in parentheses.*p<0.10, **p<0.05, ***p<0.01.

Overall, our results highlight the existence of nonlinearities in the fiscal response among the sampled countries. Though the evidence for the fiscal aggregates – primary balance and fiscal effort are at best suggestive, we find more convincing evidence for the fiscal components. In particular each fiscal component demonstrates a kinked relationship with debt-to-GDP at 50 percent. In particular countries rely on expenditure cuts when debt is below 50 percent and revenue increases when debt inches above 50 percent. Use of spending cuts as a first line of attack is not surprising if governments have ‘fat to cut’ and if raising taxes may be less popular politically. For example, as an initial policy response governments may prefer to cut spending if tax increases may prove to be more salient to citizens and could jeopardize its chances of reelection. There are also economic efficiency arguments that can explain the delayed use of tax increases until debt ratios have situated at relatively high levels as pointed out by previous studies (Krogstrup, 2002; Alesina & Ardagna, 2009). Further, distributional issues associated with tax increases in developing countries may influence the fiscal response dynamics. For example, Peralta-Alva et al. (2016) suggests that increasing the VAT – which is a huge revenue earner for many developing countries – can be extremely regressive and can lead to increases in poverty and inequality in developing countries. To the extent that these costs are significant, governments may choose to delay tax increases, opting instead to cut spending.

Heterogeneous Effects

We split the sample into high / upper-middle income and low / lower-middle income countries to examine potential differences in the fiscal reactions across country groupings.²³

Previous research (Gavin & Perotti, 2003; Reinhart, Rogoff & Savastano, 2003; Flood &

²³ Ghana, Ethiopia, Honduras and Suriname are beneficiaries under the Highly Indebted Poor Countries (HIPC) initiative. If debt relief significantly influences fiscal policy in these counties then this could potentially bias the results. Notably however our results are robust to the exclusion of these countries from the sample.

Marion, 2006) advance that relative to industrial countries, developing countries face less friendly credit market access opportunities that may exert significant influence on their fiscal policy stance. For example, developing countries may have to place much more weight on debt stabilizing fiscal policy compared to more advanced economies. If low income developing countries face less favorable credit market access opportunities relative to higher income developing countries, then fiscal policy in the former may be more sensitive to debt stabilization considerations. However, if there are more severe revenue mobilization challenges and if it is more difficult to cut non debt spending in lower income developing countries, then these ‘fiscal frictions’ might mitigate their ability to respond to debt.²⁴

Results for low / lower-middle income countries are presented in Table 6. We find that lag debt-to-GDP is positively correlated with the primary balance (0.026) and fiscal effort (0.014), but only statistically significant for the latter. Coefficient estimates for the total revenue, tax revenue and primary spending are not found to be significantly correlated with lag debt-to-GDP in low / lower-middle income countries.

²⁴ The tax/GDP ratio for low and lower – middle income countries in the sample is roughly 5 percentage points less than middle income countries.

Table 6: Results for the Fiscal Response for Low and Lower Middle Income Countries

	Primary Balance DGMM (1)	Fiscal Effort DGMM (2)	Total Revenue DGMM (3)	Tax Revenue DGMM (4)	Primary Spending DGMM (5)
lag Dependent Variable	0.597*** (0.111)	0.729*** (0.0816)	0.648*** (0.101)	0.430*** (0.155)	0.683*** (0.0725)
Lag Debt/GDP	0.0265 (0.0178)	0.0142** (0.00581)	0.0472 (0.0302)	-0.0138 (0.0148)	-0.00952 (0.0186)
Output Gap	0.00822 (0.0350)	-0.0177 (0.0237)	0.0677 (0.0451)	-0.00208 (0.0299)	0.0586* (0.0346)
Expenditure Gap	-0.0264** (0.0116)	-0.0129 (0.00996)	-0.000344 (0.0209)	-0.00492 (0.00692)	0.00499 (0.0180)
Inflation	-0.0306** (0.0149)	7.77e-05 (0.0139)	-0.108*** (0.0314)	-0.0551*** (0.0150)	-0.0680*** (0.0190)
Trade Openness	0.0361* (0.0190)	0.00107 (0.0151)	0.0803*** (0.0209)	0.0597*** (0.0109)	0.0497*** (0.0169)
Share of Agriculture in GDP	0.0490 (0.0652)	0.105* (0.0575)	-0.0385 (0.0636)	0.0736 (0.0522)	-0.105 (0.0651)
Share - Income Tax	-0.0140 (0.0246)	0.0527* (0.0270)	-0.0426 (0.0453)	0.0160 (0.0230)	-0.0360* (0.0187)
Polity IV	-0.00133*** (0.000434)	-0.000664** (0.000313)	-0.000607* (0.000350)	9.60e-05 (0.000270)	0.000789** (0.000402)
Constant					
Observations	202	202	220	219	202
No. Countries	17	17	17	17	17
AR(2) (p-value)	0.671	0.682	0.717	0.872	0.794
Hansen J (p-value)	1	1	1	1	1
Year Dummies	yes	yes	yes	yes	yes

Note: Dependent variables are expressed as a share of GDP. Results are for difference GMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995). To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables. Arellano-Bond AR (2) diagnostic tests suggest no second order serial correlation in the errors. Hansen J (p-values) are implausibly large suggesting weakened diagnostics from instrument proliferation. In regression that exclude year dummies the Hansen J (p-values) are in normal ranges and suggest that our instruments are collectively valid. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Results for upper-middle income countries are presented in Table 7 and provide an interesting contrast to that found in the case of low / low-middle income countries. Lag debt-to-GDP is positively correlated with the primary balance (0.030) and fiscal effort (0.015), but is statistically significant for the former only. Additionally, the results indicate that high / upper-

middle income countries respond to debt by making significant adjustments in total revenue (0.024) and tax revenue (0.017) but not in primary spending.²⁵

Table 7: Results for the Fiscal Response for High and Upper Middle Income Countries

	Primary Balance DGMM (1)	Fiscal Effort DGMM (2)	Total Revenue DGMM (3)	Tax Revenue DGMM (4)	Primary Spending DGMM (5)
Lag Dependent Variable	0.477*** (0.0680)	0.351* (0.188)	0.410** (0.169)	0.591*** (0.103)	0.520*** (0.127)
Lag Debt/GDP	0.0298*** (0.00700)	0.0155 (0.0124)	0.0235*** (0.00484)	0.0169*** (0.00391)	-0.00771 (0.00732)
Output Gap	0.0749** (0.0303)	0.100** (0.0416)	-0.00842 (0.0417)	0.0133 (0.0267)	-0.0750** (0.0346)
Expenditure Gap	-0.0261 (0.0266)	-0.0444* (0.0255)	0.0230 (0.0317)	0.0111 (0.0309)	0.0445* (0.0269)
Inflation	0.000894*** (0.000281)	-0.00126*** (0.000477)	0.000936* (0.000518)	-0.000889** (0.000377)	1.09e-05 (0.000357)
Trade Openness	-0.00923 (0.00919)	-0.000774 (0.0167)	-0.00791 (0.00954)	0.00130 (0.00726)	-0.00100 (0.0119)
Share of Agriculture in GDP	0.0366 (0.0776)	0.131 (0.0925)	-0.103 (0.0706)	-0.0356 (0.0431)	-0.126** (0.0587)
Share - Income Tax	-0.0171 (0.0213)	0.0489 (0.0415)	-0.0305 (0.0351)	0.0179 (0.0190)	-0.0178 (0.0353)
Polity IV	-0.000719 (0.000522)	-0.000215 (0.000444)	-0.00131 (0.000822)	-0.000597 (0.000534)	-0.000426 (0.000445)
Constant					
Observations	386	386	392	392	386
No. Countries	26	26	26	26	26
AR(2) (p-value)	0.244	0.078	0.810	0.970	0.963
Hansen J (p-value)	1	1	1	1	1
Year Dummies	yes	yes	yes	yes	yes

Notes: Dependent variables are expressed as a share of GDP. Results are for difference GMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995). Relative to first difference transformations, FOD minimizes data loss in unbalanced panels; and since lagged observations are not used in the transformation they are valid instruments. We treat the lag dependent variable and the lag debt-to-GDP as potentially endogenous. Hansen J (p-values) are implausibly large suggesting weakened diagnostics from instrument proliferation. In regression that exclude year dummies the Hansen J (p-values) are in normal ranges and suggest that our instruments are collectively valid. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

The results suggest that both low / lower-middle income and high / upper middle income countries are fiscally responsible, with positive and statistically significant adjustments in the fiscal effort or primary balance. However, whereas low / lower-middle income countries do not

²⁵ The Hansen J p-values in all the country subsample regressions are inflated due to instrument proliferation. As a result we are unable to definitively test the validity of the instruments for the sub-sample analyses, particularly in the regressions for low and lower-middle countries.

significantly adjust on the revenue or spending margins, high / upper-middle income countries adjust mainly by increasing revenues. These results might suggest greater ability on the part of high / upper middle income countries to increase revenues in response to growing debt, and perhaps an inability on the part of low / lower-middle income countries to do the same.

Sensitivity Analysis

We test the robustness of the baseline results using alternative specifications of the DGMM regression model. To do this we simply add covariates to the model and also experiment with different combinations of covariates to see if this would significantly change the coefficient estimates from the baseline results. We systematically add the following controls; net ODI/GNI, a dummy for whether the country had a fiscal rule and a dummy for whether the country had an IMF agreement in place in regressions for each fiscal outcome. The results presented in Tables A.3 – A.7 of Appendix A suggest that the baseline estimates are robust to the inclusion of the additional covariates.

We also vary the sample over time to check whether the baseline results are robust across different time periods. First, we focus on observations between 1990 and 2008 to exclude years post the great recession. Though developing countries are argued to have been somewhat ‘sheltered’ from the major effects from the great recession, it is likely that creditors would have become much more cautious thereby influencing countries ability to access credit. Additionally, the governments in developing countries themselves may have assessed the potential risks from the recession and implemented policies to mitigate same. Excluding these years could potentially

avoid confounding from effects related to the great recession. Additionally, we examine observations over a shorter time period between 1990 and 2000.²⁶

Results for the primary balance and fiscal effort for the baseline (1990 – 2011) and the two sub-periods are presented in Table A.8 of Appendix A. In general, the results for both outcomes are robust to the contraction of the sample period. The estimated coefficients, however, get progressively larger as we move from the longest sample period to the shortest sample period. Table A.9 present results for total revenues, tax revenues and primary spending. These results are also robust to contraction of the sample period and we find similar patterns in the size of the estimated coefficients across the various sub periods.

Conclusion

This essay makes several contributions to a literature that examines countries fiscal response to debt. Firstly, our analysis focuses on developing countries – a group that had previously received scant attention in the literature. Secondly, the few studies that examine the fiscal reaction function for developing countries did so over a relatively limited time span. Our analysis utilized an expanded time series relative these earlier studies. We also point out a serious omission in the literature – by emphasizing the importance of including ‘fiscal effort’ as a key indicator in debt sustainability analysis. We are also the first to decompose key components of the primary balance and examine the specific margins of response for developing countries.

We find that, in general, developing countries are fiscally responsible as evidenced by the positive and significant response in the primary balance and ‘fiscal effort’ to debt-to-GDP. For

²⁶ We also test for long run effects using a distributed lagged model but found no significance beyond the first period lag. However the results from this model may be dubious because of multicollinearity of the lagged debt ratios and thus we do not report the results from these regressions.

the full sample we find evidence that countries adjust along both the revenue and expenditure margins at roughly the same rate. This sort of measured adjustment might be indicative of deliberate policy on the part of governments, not to engage in asymmetric policy changes that could impact negatively on overall economic welfare. We also find evidence of nonlinearities, particularly for the fiscal components, with larger increases in revenues when debt is relatively high and larger reductions in primary spending when debt is relatively low. The pattern of responses across the various margins suggests that countries delay tax or general revenue increases until debt is relatively high and cut spending when debt is relatively low. The adjustment decision might be driven by several factors. For example, political economy arguments point to the relative unpopularity of tax increases compared to some expenditure cuts. Countries might therefore be encouraged to use spending cuts as their first ‘line of attack’ and resort to tax increases when the debt problem becomes more acute. There are also economic arguments – taxes may be more distortionary than spending cuts and is thus relegated as a secondary option to stabilize debt.

We find that both low / lower-middle income and high / upper-middle income countries have been fiscally responsible – with positive correlations between debt and fiscal effort primary surplus, respectively. However, high / upper-middle income countries demonstrate a greater propensity to adjust along the revenue and tax revenue margins highlighting an important difference in the fiscal reaction across the two groups of countries. The relatively strong response in general revenues and tax revenues for high / upper middle income countries and the lack thereof for low / lower-middle income countries might be indicative of systemic revenue mobilization challenges facing the latter. One obvious implication of this is that these countries

must strengthen their revenue generating capacity in order to better sustain their debt. This is particularly true for countries with relatively high levels of debt.

Essay 2: The Effect of Taxpayer Service Provision on Tax Compliance for Large Taxpayers in Jamaica

Introduction

The fiscal landscape across developing countries is characterized by significant revenue mobilization challenges, stemming from *inter alia* critically low levels of tax compliance. In exploring the potential causes of low tax compliance in developing and transition economies, Alm and Martinez – Vazquez (2003) attribute much of the effects to weak fiscal institutions. They stress that the role of tax administration surpasses merely securing revenues for the state and must also include ensuring taxpayer satisfaction, equity and social welfare. These arguments are in line with more recent views that highlight the need for a balance between enforcement and more facilitatory approaches, grounded in a commitment to an implicit ‘psychological’ contract between the taxpayer and the tax administration (Feld & Frey, 2002; Kirchler, Hoelzl & Wahl, 2008).²⁷

The thrust of recent tax administration reforms across countries aim at increasing both enforced and voluntary compliance within the context of an overarching risk based approach. In particular, tax administrations recognize the importance of tax morale or non-pecuniary factors in determining taxpayer compliance and have sought to complement traditional enforcement strategies with a softer approach – such as the provision taxpayer services. Additionally, to mitigate risk exposure, tax administrations have sought to adopt a more strategic focus on individuals and firms that pose the greatest risk to revenue collections – example large taxpayers.

²⁷ Kirchler (2007) argues that the cops and robbers approach to tax administration fuels distrust and adversarial tendencies between the tax administration and the taxpayer, but a service - client approach can encourage greater cooperation and improve voluntary compliance.

This essay examines the effect of the provision of taxpayer services – through the Large Taxpayer Office (LTO), on the timeliness and completeness of filing and payment of the corporate income tax (CIT) and general consumption tax (GCT) for large taxpayers in Jamaica. We focus on the taxpayer’s decision to file and pay taxes conditional on reporting positive tax liabilities. These are important margins of response particularly in developing countries with relatively weak tax administrations that find it difficult to collect reported taxes. Moreover, where tax administration resources are limited, and where the outcomes of expensive audits are uncertain, it may be more prudent to focus on collecting pledged taxes.

The empirical strategy exploits quasi-experimental variation in the intensity of service delivery for taxpayers selected into the LTO. A key criterion for selection into the LTO is having gross receipts greater than or equal to an arbitrary threshold of J\$500 million (US\$5.7 million).²⁸ We use a regression discontinuity design (RDD) that exploits the discrete jump in the probability of selection at the threshold. This approach compares the compliance behavior of those located just to the right of the threshold – who are selected into the LTO; to otherwise similar taxpayers located just the left of the threshold – who are marginally not selected into the LTO. Assuming all other key taxpayer characteristics transition smoothly across the threshold, the RDD estimates are causal.

This essay relates to a growing literature that use experimental and quasi-experimental approaches to examine the effect of enforcement and tax morale factors on tax compliance for developing countries in the Latin America and Caribbean (LAC) region (Carillo, Pomeranz, & Singhal, 2014; Ortega & Scartascini, 2015; Castro & Scartascini, 2013; Pomeranz, 2013). This is the first for an English speaking Caribbean country. This paper also links closely to literature that

²⁸ The US dollar conversion was done using a 4 year (2009 – 2012) average annual exchange rate of \$87.35 JMD to \$1 USD using data from the Bank of Jamaica.

examines taxpayers' behavioral response to discontinuities in the tax structure. Whereas some studies examine behavioral responses to notches or kinks in the tax rate structure (Saez, 2010; Keleven & Waseem 2012), others leverage exogenous changes in monitoring and enforcement intensity that occur around some revenue threshold (Almunia & Rodriguez, 2014; Sanchez, 2013). This paper relates most closely with this second strand of literature, but instead of an 'enforcement notch' we examine compliance response to changes in the intensity of taxpayer service delivery.

First evidence of the compliance effects of taxpayer services come from a handful of lab and field experiments (Alm, Chetty, Jone & McKee, 2011; Mckee, Siladke & Vossler, 2011; Vossler & Mckee, 2013; Kosonen & Ropponen, 2015) and focus on taxpayer information services. These studies find generally positive effects on filing and reporting compliance but also report limits on the impact of taxpayer information services. Mckee et al. (2011) find that post audit, subjects are less likely to request information assistance and are more likely to evade. Kosonen and Ropponen (2013) find that taxpayer information services are ineffective in boosting compliance for relatively more complex aspects of the VAT law in Finland. In this paper we go beyond examining the effect of taxpayer information services and examine the effects of a range of taxpayer services. We also explore the effects of these services, conditional on the strength of the legal enforcement framework of the taxing regime, to test the relationship between service delivery and enforcement strength.

We find that taxpayer service provision did not significantly improve filing or payment compliance for the CIT, but had generally positive compliance effects for the GCT. Taxpayers are (22 pp) less likely to file GCT returns late and also reduce the number of days late that returns are filed by about 14 days. Taxpayers are also (17 pp) less likely to pay GCT late and

reduce the number of days late that payments are made by about 247 days. The amount of GCT and the share of reported GCT paid on time increases by about J\$4 million and 27 percent respectively. We attribute the null effects for the CIT and positive compliance effects for the GCT to the relatively stronger legal enforcement framework of the latter. We take this as suggestive evidence of a complementarity in the relationship between the strength of the legal enforcement framework of the taxing regime and the provision of taxpayer services. We also find heterogeneous effects in the case of the GCT, with positive compliance response for non-financial sector taxpayers and non-importers but null effects for financial sector taxpayers and importers. We attribute this to possible substitutability between taxpayer service provision and external enforcement mechanisms such as industry oversight bodies and tax compliance requirements to receive import licenses.

The results highlight a possible limitation in the use of taxpayer services to drive compliance in the absence of a robust legal enforcement framework. One implication of this finding is that tax administration and tax policy reforms must be balanced in its focus on enforcement and the ‘softer approach’ in order to improve compliance among large taxpayers. The results also point to scope for the use of external enforcement mechanisms to substitute for expensive tax administration interventions such as service provision.

Background

The Corporate Income Tax and the General Consumption Tax

The CIT rate for tax years 2009 – 2012 was 33.33 percent and applied to firms' reported profits. The GCT is a value added tax (VAT) and was applied at standard rates of 16.5 and 17.5 percent over the sample period.²⁹ The CIT and GCT are huge contributors to the overall tax revenue, accounting for about 11 and 17 percent to total tax collections respectively.³⁰ Jamaica's tax administration machinery has long been criticized as weak and inefficient. An obvious consequence of this is low tax compliance. In 2011 filing and payment compliance for the CIT – measured as whether the taxpayer filed and paid on time – were 40 and 53 percent respectively. In the same year, for the GCT, 83 percent of taxpayers filed on time and 88 percent paid on time.

One likely explanation for the sizeable difference in compliance rates for the CIT relative to the GCT is the stronger legal enforcement framework of the latter. The penalty structure across the two taxes presents an interesting contrast. Failure to file CIT on time attracts a fine of J\$5,000 (US\$57) and interest of 40 percent per annum is charged against outstanding CIT liability. On the other hand, failure to file GCT on time attracts a fine of J\$2,000 (US\$23) or 15 percent of the tax due and payable, whichever is larger. Interest of 2.5 percent compounded monthly is charged against the sum of outstanding GCT liabilities, penalties and surcharges. For large taxpayers, this structure implies higher penalties and interest for delinquents under the GCT relative to the CIT, with the difference between the two increasing in the amount of unpaid tax liabilities as shown in Figure B.1 of Appendix B. One obvious implication of this is that (larger) taxpayers have a clear incentive to be more compliant with the GCT than the CIT as evidenced

²⁹ From May 2005 – December 2009 the standard GCT rate was 16.5 percent. The standard GCT rate was temporarily increased to 17.5 percent in January 2010 before being lowered to 16.5 percent in June 2012.

³⁰ In this research we focus on the local and not the international component of GCT. GCT's total (local and international) contribution to tax collections averaged about 31 percent over the last six fiscal years.

by the compliance patterns in Figure B.2 in Appendix B. Panel A shows much lower rates of late filing and payment for the GCT relative to the CIT. Panel B shows larger amounts of CIT paid on time but a larger share of reported GCT paid on time. The compliance patterns across taxes appear to be driven largely by key features of the legal enforcement framework. Comparing the compliance effects of the LTO across the CIT and GCT provides an opportunity to examine the potential effect of taxpayer service provision in a weak and strong enforcement context respectively.

The Large Taxpayer Office

Like most developing countries, the tax regime in Jamaica is characterized by extremely high dependence on a few large taxpayers for revenue. In 2011 the top 1 percent of taxpayers accounted for 82 percent and 66 percent of reported CIT and GCT respectively.³¹ This dependence exposes the government to an extremely high level of risk from non-compliance of few taxpayers, and provides impetus for the tax administration to boost compliance efforts in general, but particularly for large taxpayers. To manage this risk, TAJ established the LTO in April 2009. Taxpayers are assigned to the LTO if they meet any of three criteria. The first and principal criterion is if annual gross receipts are greater than or equal to J\$500 million (US\$5.7 million). Secondly, if the total annual taxes paid is greater than or equal to J\$50 million.³² Thirdly, if related to a primary LTO client through for example common ownership, a subsidiary or branch.

³¹ We use reported taxes instead of actual tax payments because the payment data provided by TAJ does not adequately identify specific payment components and therefore potentially comingles principal tax payments with penalties and interest.

³² In general taxpayers who pay at least \$50 million in taxes are large and most likely will gross more than \$500 million in sales / gross receipts annually.

An important feature of the LTO is its central focus on service provision. Its stated mission is “*to promote voluntary compliance and enhance revenue collection by providing exemplary specialized service to the large taxpayer population through a team of highly motivated and results oriented professionals*”. At the core of the LTO operations are client relationship managers (CRM). These positions were specially created to channel specialized services to large taxpayers. Once selected to the LTO each taxpayer is assigned a CRM who will serve as the main point of contact with the tax administration. The CRM will introduce and provide or facilitate the provision of the range of services offered by the LTO. These services include but are not limited to, the provision of tax advice, processing of tax compliance certificates, stamping of documents, tax seminars and workshops, filing and payment reminders, filing and payment facilitation, registration and reconciliation of tax accounts.³³ Although some of these services are available to the general taxpayer population at the various tax offices or through the customer care center, we argue that there is a significant increase in the intensity and efficiency of service delivery for LTO taxpayers.³⁴ Moreover the ‘one-on-one’ between LTO clients and CRM guarantees more efficient and convenient service delivery.

Theory

The standard economic model of tax evasion is outlined by Allingham and Sandmo (1972) (AS) – an adaptation of Becker’s (1968) economics of crime approach – and models individuals’ tax compliance behavior as a rational decision. Their model predicts that higher levels of evasion are associated with low detection probabilities and low pecuniary costs if

³³ CRM’s also help to facilitate audits carried out by TAJ’s auditors. Their main role here is to ensure that the audit process is smooth and not overly disruptive for the taxpayer’s normal business operations.

³⁴ Interviews with CRM’s from the LTO confirm that the level of services offered to LTO clients is superior to that available regularly through the tax offices and call centers. They also point out that there are clear benefits to the taxpayers from the one-on-one interaction with the CRM which are derived not only through access to services but also the development of good relations.

caught. The theoretical predictions of the AS model however have been criticized as inadequate in explaining observed patterns of tax compliance across countries. Complementary theories of tax compliance attempt to overcome some of the shortcomings of the standard approach and posit that non-pecuniary or tax morale factors play an important role in explaining taxpayer behavior (Luttmer & Singhal, 2014).

Whereas the standard theory models taxpayers reporting decision, more recent work examine taxpayers payment decision conditional on income being declared, or on established tax liabilities (Hallsworth, List, Metcalfe & Valev, 2014; Perez-Truglia & Troiano, 2015), a question much more closely related to this research. We adopt a simple model of payment compliance advanced by Hallsworth et al. (2014). We assume a two period model where taxpayers earn income (Y_i^G) – drawn from an i.i.d. probability distribution $f(Y_i^G)$ – in period one and none in period two.³⁵ The taxpayer is required to pay taxes on reported income at some rate (t) and the tax liability is given by:³⁶

$$T_i = t * \max[0, Y_i^G] \quad (8)$$

We take reported income Y_i^G as given, but the taxpayer must decide whether to pay taxes in period one ($x=1$) or period two ($x=2$). We model the taxpayer's payment decision as a function of the real interest rate, compliance costs and a moral or psychic cost from non-compliance. The payment decision is made to maximize after tax income (Y_i^N) based on the following:

³⁵ We take reported income as given and our model focuses solely on the decision to pay, conditional on reporting some positive tax liability.

³⁶ The model assumes that the decisions to file and pay taxes are made simultaneously. TAJ confirms that taxpayers typically file and pay (at least some portion) taxes at the same time.

$$Y_i^N = \begin{cases} Y_i^G - T_i - C_i & \text{if } x = 1 \text{ and } Y_i^G - T_i - C_i \geq 0 \\ Y_i^G + (1 + r)T_i - M_i - (1 + \alpha)T_i & \text{if } x = 2 \text{ or } Y_i^G - T_i - C_i < 0 \end{cases} \quad (9)$$

From equation (9), taxpayer (*i*) faces compliance cost ($C_i > 0$) in the first period but not in the second. This is a simplifying but rationalize-able assumption. For example it is plausible that compliance costs in period one are higher simply because taxpayers have to expend considerably more effort to complete filing and payment by a stipulated deadline.³⁷ The compliance cost in period one is juxtaposed against the costs and benefits from delaying payment until period two. Delaying payment means that the taxpayer can benefit from interest earned (*r*) on taxes not paid in period one. However the taxpayer also faces a penalty of α , that is proportional to the amount of tax owed and which takes values $[0, 1]$. Lastly, we assume that non-compliance in period one is associated with a moral cost ($M_i > 0$) incurred in period two, and is induced by the provision of taxpayer services.

The taxpayer will pay in period one if the value of doing so exceeds that of delaying until period two. The payment decision is therefore captured by the following condition:

$$Y_i^G - T_i - C_i > Y_i^G + (1 + r)T_i - M_i - (1 + \alpha)T_i \quad (10)$$

From equation (10), the compliance effects of the LTO come through several channels. The first is through lower compliance costs. If taxpayer services that clarify complex or ambiguous tax laws and that facilitate speedy and convenient filing and payment of taxes can significantly reduce compliance costs (pecuniary and non-pecuniary) in period one, taxpayers may be encouraged to pay on time. The second is through a behavioral channel. In addition to

³⁷ Complexities in the tax law can impose huge compliance costs – in terms of time and money – as taxpayers struggle to accurately calculate their taxes in order to file and pay by the due date. The numerous requests for filing extensions received by the TAJ may be an indication of the relatively large compliance costs incurred when trying to comply with the stipulated filing deadline. In the second period compliance costs are arguably lower since there is no longer an effective time deadline by which to comply in order to avoid penalties.

the services provided, closer interaction between the tax administration and taxpayers within a customer centric context can improve taxpayers' perception of the tax administration. This can increase the moral or psychic costs of delaying payment – crowd in tax morale – and induce higher levels of (voluntary) compliance (Feld & Frey, 2002). Notwithstanding the predictions of the theoretical model, the impact on compliance is an open empirical question.

Data and Empirical Strategy

Data

We use administrative data for the CIT and the GCT for 2009-2012.³⁸ We restrict our baseline sample to taxpayers with reported gross receipts of between J\$100 million (US\$ 1.1 million) and J\$1 billion (US\$11.4 million) who reported positive tax liabilities. The resulting sample sizes are 2,432 for the CIT and 34,764 for the GCT. From tax returns, we collect data on filing date, gross receipts and reported tax liabilities. From the payments data we collect information on the payment date and the amount paid. Using these data, we construct our outcome measures for filing and payment compliance. For filing compliance we examine two outcomes. The first is captured by a dichotomous variable set equal to 1 if a taxpayer filed late and 0 if filed on time.³⁹ The second is the number of days a return is late. Essentially these measures together capture filing compliance along an extensive margin – whether a taxpayer filed late or not; and an intensive margin – how late was the taxpayer in filing.

³⁸ We examine compliance responses for taxpayers (firms) that file and pay the “company income tax final return” on form ITO2. We do not consider corporate tax returns filed on forms ITO3 and ITO4 which are the designated forms for “unincorporated bodies other than life assurance companies” and “life assurance companies” respectively. For the GCT we focus on the standard GCT return and do not include the quick return or returns for special tourism activities.

³⁹ For the CIT, taxpayers are required to file their final or annual return by March 15 following the year of assessment. If this date falls on a weekend, we record as late, if the taxpayer filed after the first business day of the following week. This is consistent with the practice of the tax administration.

For payment compliance we examine two sets of outcomes - in total four outcome variables. The first two captures the timeliness of payments and are measured the same way we measured filing compliance above. In the second set we examine two additional payment outcomes - the amount of taxes paid on time and the share of reported taxes paid on time. Because the payment data gives aggregate amounts received from taxpayers, we are unable to identify specific payment components, i.e. whether amounts paid are solely taxes or a combination of taxes and other charges and penalties. By restricting the analysis to payments made on or before the due date – we are arguably better able to capture taxpayers ‘real’ compliance response since payments made after the due date with respect to a particular filing period are more likely to include amounts for penalties, interest and audit assessments. As such, the amount of taxes paid on time and the share of reported taxes paid on time are arguably better measures of taxpayers’ response to the services provided through the LTO, in relation to their contemporaneous tax liabilities.

We identify LTO clients using the client listing provided by TAJ. Treatment is indicated using a dichotomous variable set equal to 1 if the taxpayer appears on the client list and 0 otherwise.⁴⁰ Other data on taxpayer characteristic such as age, number of employees, economic sector and other financial data are from the tax returns and the tax registry databases. Tables 8(a) and 8(b) provide summary statistics for data used for the CIT and GCT analysis respectively.

⁴⁰ The client listing provided was for taxpayers in the LTO as at 2013. We were unable to get a list for each of the years 2009 – 2012 from TAJ and so we are not able to tell when a particular taxpayer came onto the register. The analysis therefore assumes taxpayers listed on the client register in 2013 were clients for each of the years used in the analyses. If the LTO was effective in encouraging compliance then this could potentially bias our estimates downward. We re-estimate the model using only those taxpayers listed as LTO parent companies as our treatment group since it is most likely that the larger parent companies would have been on the client listing since its inception. The results are qualitatively similar to the baseline, with slightly larger coefficients.

Table 8: Summary Statistics, CIT & GCT

Table 8 (a): Summary Statistics for the CIT

Variable	Observations	Mean	Std. Dev.
LTO (Treatment)	2,432	0.17	0.37
Gross receipts annual (J\$ Millions)	2,432	307.00	209.00
Filed late	2,394	0.37	0.48
No. days filed late	2,394	38.59	96.79
Paid late	1,800	0.54	0.50
No. days paid late	1,800	289.21	468.69
CIT Paid (J\$ millions)	1,800	3.57	13.90
Share CIT paid on time	1,800	0.45	0.53
Estimated CIT Paid (J\$ millions)	2,432	2.76	11.00
CIT Reported (J\$ millions)	2,432	3.26	8.26
Financial Sector	2,432	0.05	0.21
Number of Employees	2,124	48.34	90.51
Age	2,432	10.55	4.40

Notes: Tax return data are for taxpayers with reported gross receipts between J\$100 million and J\$1 billion who report positive tax liabilities for tax years 2009 – 2012. Payment data are for tax years 2009 – 2011 as data for 2012 were unavailable.

Table 8 (b): Summary Statistics for the GCT

Variable	Obs	Mean	Std. Dev.
LTO (Treatment)	34,764	0.22	0.42
Gross receipts annual (J\$ Millions)	34,764	296	203
Filed late	34,739	0.06	0.23
No. days filed late	34,739	3.76	35.32
Paid late	34,764	0.10	0.30
No. days paid late	34,764	78.73	349.59
GCT Paid (J\$ Millions)	34,764	1.22588	2.73232
Share GCT paid on time	34,762	0.90	0.74
Tax Arrears (J\$ millions)	34,734	-0.009	0.775
Financial Sector	34,764	0.07	0.26
Age	33,951	8.72	1.98

Notes: Data are for taxpayers with reported gross receipts between J\$100 million and J\$1 billion, who report positive tax liabilities between May 2009 and December 31, 2012. We capture data starting roughly two months after the establishment of the LTO in April 2009.

Empirical Analysis

The empirical strategy adopts a fuzzy regression discontinuity design (RDD) that exploits an exogenous jump in the intensity of taxpayer service delivery that occurs when a taxpayer reaches gross receipts of J\$500 million (US\$5.7 million) and is selected into the large taxpayer office (LTO). Despite being the primary criteria, as explained above, the gross receipts threshold is not the sole selection criteria. Together with the other criteria and the operationalization of the selection process, selection into the LTO is non-deterministic at the threshold. Rather the probability of being selected (treated) is represented by:

$$P[T_i = 1 | S_i] \begin{cases} g_1(S_i) & \text{if } S_i \geq S_0 \\ g_0(S_i) & \text{if } S_i < S_0 \end{cases} \text{ where } g_1(S_i) \neq g_0(S_i) \quad (11)$$

The running variable S_i and the probability of treatment are related as follows:

$$P[T_{i,t} = 1 | S_i] = g_0(S_i) + [g_1(S_i) - g_0(S_i)]D_i \quad (12)$$

Where $D_i = 1(S_i \geq S_0)$

Estimation of the local average treatment effect (LATE) in a fuzzy RDD can be modeled by a two stage (2SLS) approach as set out in the following model:

First Stage:

$$E [T_i = 1] = \alpha_0 + \alpha_1 D_{i,t} + \alpha_2 f(s_{i,t}) + \alpha_3 f(s_{i,t}) * D_{i,t} + \mu_{i,t} \quad (13)$$

Where the running variable ($s_{i,t}$), is the gross receipts of taxpayer i at time t , centered at the threshold value.⁴¹

⁴¹ We follow the literature and center gross receipts around the threshold [$S_i = (S_i - \text{J\$500M})$]. This ensures that the treatment effect at the threshold can be read from the coefficient on the treatment indicator in models that include interaction terms.

Second Stage:

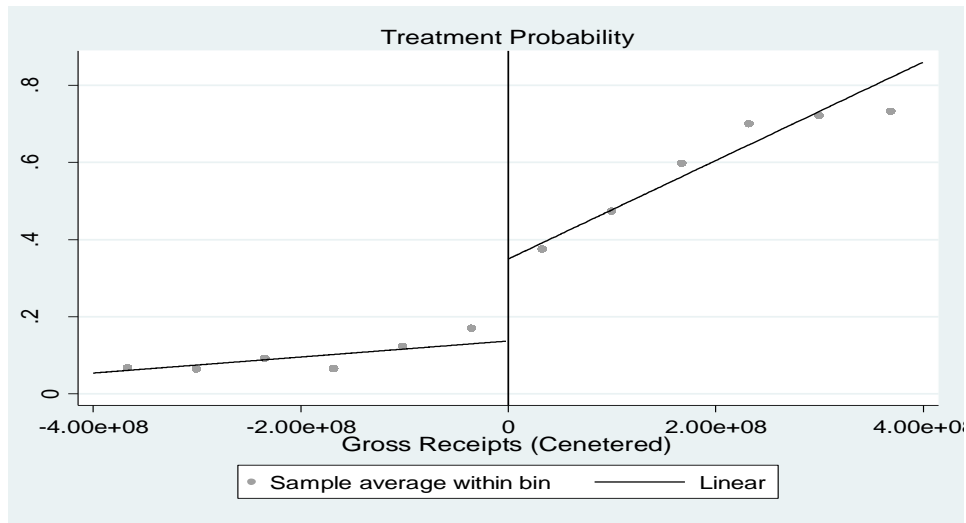
$$Y_{i,t} = \beta_0 + \beta_1 T_{i,t} + \beta_2 f(\widetilde{s}_{i,t}) + \beta_3 f(\widetilde{s}_{i,t}) * T_{i,t} + \varepsilon_{i,t} \quad (14)$$

The first stage regression estimates the probability of treatment using the firm's location relative to the gross receipts threshold $D_{i,t}$ and the interaction with gross receipts as instruments. $f(\widetilde{s}_{i,t})$ is a flexible polynomial function to account for possible nonlinearities in the relationship between the outcome variables and the running variable. Our baseline model adopts a linear specification. The second stage uses the fitted values from the first stage to estimate the local average treatment effect (LATE) on the dependent variable $Y_{i,t}$ – which captures the various measures of tax compliance outlined above. The LATE is given by β_1 . In alternative specifications of the model we include key taxpayer characteristics as controls in $X_{i,t}$ and year dummies τ_t to test for robustness.

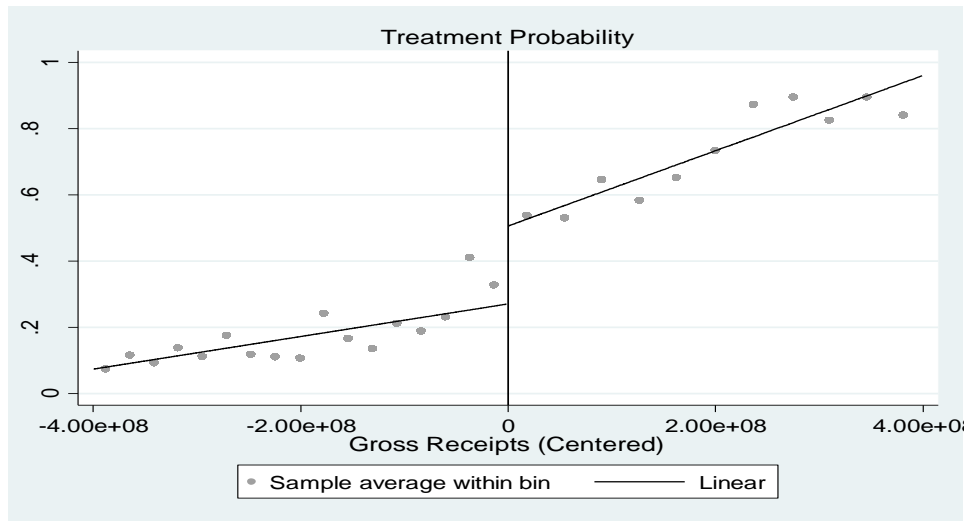
We first examine the validity of the RDD. Panels A and B in Figure 2 show regression discontinuity plots of the probability of treatment under the LTO for the CIT and GCT samples respectively. The plots fit local linear regressions on either side of the threshold with local sample averages of the outcome variable – in this case the probability of treatment, within bins of the running variable – in this case gross receipts (centered at J\$500 million). Both figures show a discontinuous increase in treatment probability at the gross receipts threshold, suggesting the RDD is valid.

Figure 2: Test of Identification - First Stage RDD

Panel (a): Discontinuity in Treatment Probability for the CIT



Panel (b): Discontinuity in Treatment Probability for the GCT



Another key identifying assumption of the RDD is that taxpayers are not able to perfectly manipulate reported gross receipts; either to locate below the threshold to avoid selection or above the threshold to invite selection. We argue that because the gross receipts threshold is not the sole factor used in assigning taxpayers – the others being whether a taxpayer is related to an LTO client and if total taxes paid is greater than or equal to J\$50 million, then manipulation by

misreporting gross receipts become less viable. If taxpayers significantly manipulate reported gross receipts in order to avoid or invite treatment we expect to see bunching around the threshold. We test for structural breaks in the density of reported gross receipts by first examining simple histogram density plots for the CIT and GCT samples in Figure B.3 in Appendix B. More formally we adopt Mcrary (2008) density manipulation test for the respective samples presented in Figure B.4. Both tests show no evidence of bunching or manipulation of the running variable.

In other tests we examine the existence of discontinuities at other plausible points along the gross receipts distribution – ‘placebo thresholds’, based on TAJ’s segmentation of the taxpayer population. Figure B.5 in Appendix B present the results for test for discontinuities at placebo thresholds at J\$100 million (US\$1.1 million) and J\$1 billion (US\$ 11.4 million). The RD plots indicate that treatment probability does not exhibit a discontinuity at either of the placebo thresholds. Further tests of the credibility of the RDD examine the transition of other covariates across the gross receipts threshold in Figures B.6 and B.7 in Appendix B. This is to assuage concerns that other factors might also be changing discontinuously around the threshold that could impact the compliance outcomes we examine. In general, the plots for both the CIT and GCT samples show taxpayer characteristics transitioning smoothly across the threshold, except for an indicator for taxpayers in the financial sector.⁴²

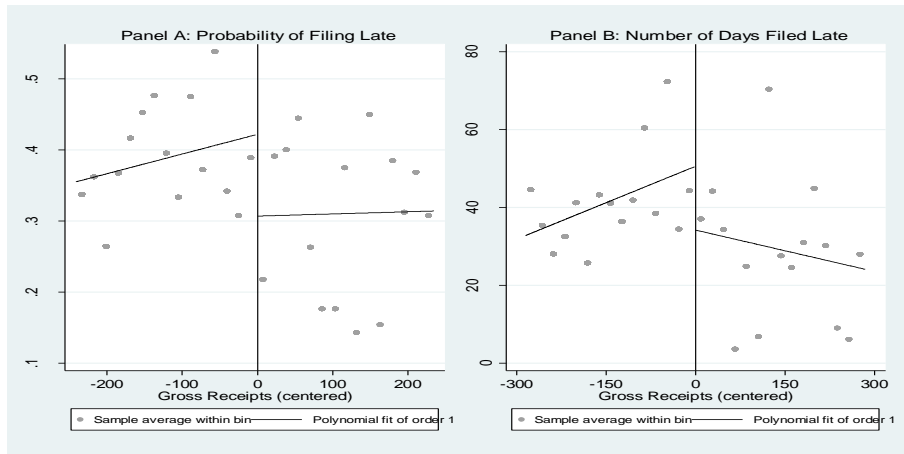
⁴² The confidence intervals on the RD plot for the financial sector indicator are large and suggest that the discontinuity observed is not statistically significant. We run regressions excluding financial sector taxpayers to test for potential bias, for both the CIT and GCT sample. The results are similar to the baseline model.

Results

Filing Compliance for the CIT

RD plots presented in Figure 3 show a reduction in the probability of late filing and the number of days late that CIT returns are filed. However the estimates are imprecise, and calls into question their statistical significance.

Figure 3: Effect of the LTO on Timely Filing of the CIT



Notes: Panel A shows RD plots for the probability of late filing and panel B shows RD plots for the number of days late that returns are filed. Bandwidths are chosen to approximate the optimal (IK) bandwidth for the respective outcomes. RD Plots done using IMSE – evenly spaced (ES) method with spacing estimators proposed by Calonico et al. (2014b). The approach fits linear regressions that approximate the conditional mean of the outcome variables to the left and right of the cut off.

Table 9 presents estimates of the LATE from 2SLS regressions, using data within ranges of gross receipts that correspond with the optimal bandwidth proposed by Imbens and Kalyanaraman (2009) (IK).⁴³ Panel A shows a reduction in the probability of late filing by about 90 percentage points (Column 1) and Panel B shows a reduction in the number of days late that

⁴³ Baseline results from 2SLS regressions using the optimal IK bandwidth are reported for all compliance outcomes examined. Results for alternative bandwidths (+/- J\$20 million) are presented in the results tables for the respective outcomes as robustness checks.

returns are filed by 190 days (column 1). However the effects are imprecisely estimated and in both cases are statistically insignificant.⁴⁴

Table 9: Effect of the LTO on Timely Filing of the CIT

	Filing Compliance					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Probability of Filing Late						
LTO	-0.904 (1.241)	-0.315 (0.835)	-0.843 (1.519)	-0.0722 (0.990)	-0.916 (1.059)	-0.435 (0.737)
Bandwidth (J\$ Millions)	240	240	220	220	260	260
Observations	938	837	823	732	1,038	927
Year Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes
Panel B - Number of Days Filed Late						
LTO	-190.2 (188.1)	-125.7 (137.0)	-209.1 (219.5)	-126.0 (152.1)	-165.0 (157.9)	-118.6 (121.9)
Bandwidth (J\$ Millions)	286	286	266	266	306	306
Observations	1,179	1,063	1,063	952	1,307	1,178
Year Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes

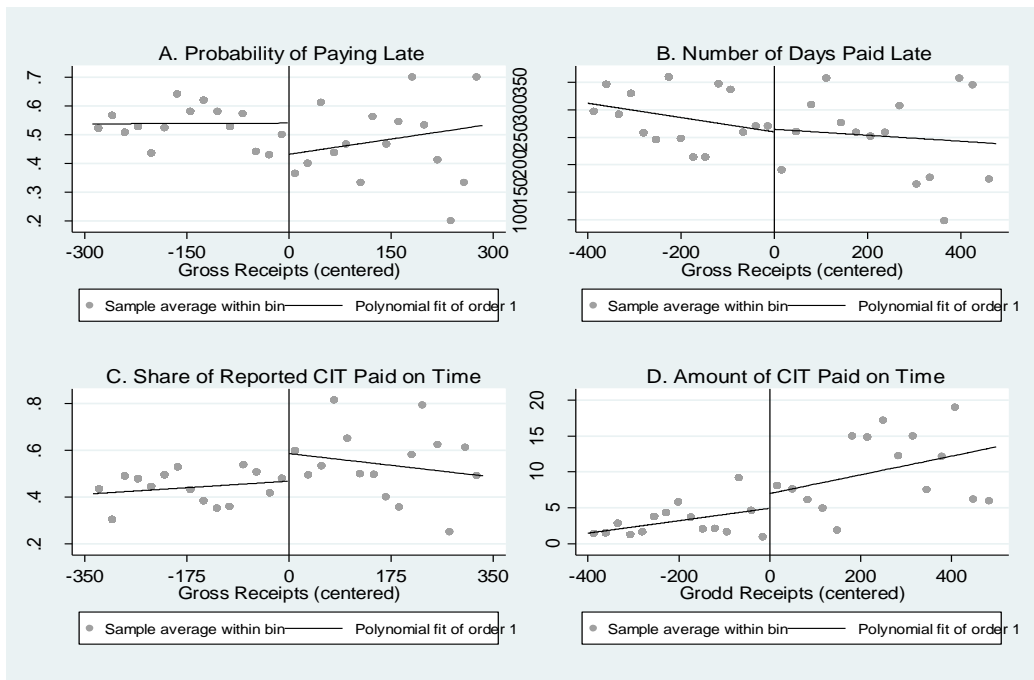
Note: This table presents estimates for filing compliance for the CIT. The dependent variable in Panel A is an indicator equal 1 if the taxpayer filed late and 0 otherwise. The dependent variable in Panel B captures the number of days after the due date that taxpayers file a CIT return. The results presented in columns 1-2 are for the optimal bandwidth as proposed by Imbens and Kalyanaraman (2009), with and without controls respectively. Columns 3-4 present results for a smaller bandwidth (IK – J\$20 million) and columns 5-6 for a larger bandwidth (IK + J\$20 million). Controls include the ‘age’ of the taxpayer, the amount of estimated CIT paid, reported CIT and a dummy that capture whether or not the taxpayer operates in the financial sector. Regressions adopt a linear specification. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

⁴⁴ We also examine the filing response for a larger sample of taxpayers, which include all taxpayers with gross receipts between J\$100 million and J\$1 billion who filed a tax return for the period of analysis, which therefore includes those that reported zero tax liabilities. The results in Table B.1 in Appendix B indicate no significant effects on filing compliance.

Payment Compliance for the CIT

We examine payment compliance for the CIT for tax years 2009 to 2011.⁴⁵ RD plots in Figure 4 show a reduction in the probability of late payment in panel A but no effect for the number of days late that CIT is paid in panel B. The share of reported CIT paid and the amount of CIT paid on time are higher just right of the threshold, as shown in panels C and D respectively.

Figure 4: Effect of the LTO on Timely Payment of the CIT



Note: Bandwidths are chosen to approximate the optimal (IK) bandwidth for the respective outcomes. RD Plots done using the default IMSE – evenly spaced (ES) method with spacing estimators Calonico et al. (2014b). The approach fits linear regression curves that approximate the conditional means of the outcome variable to the left and right of the cut off.

⁴⁵ CIT payments data for tax year 2012 were not available.

The directional effects from the RD plots are again re-enforced by the LATE estimates in Table 10, but again, are statistically insignificant. The null effects hold for alternative bandwidth choices and model specifications. Interestingly the point estimates from 2SLS are large and may suggest significant economic gains in compliance. However the lack of statistical significance does not allow for any clear determination of the effects of the LTO on CIT compliance outcomes examined.

Table 10: Effect of the LTO on Payment Compliance for the CIT

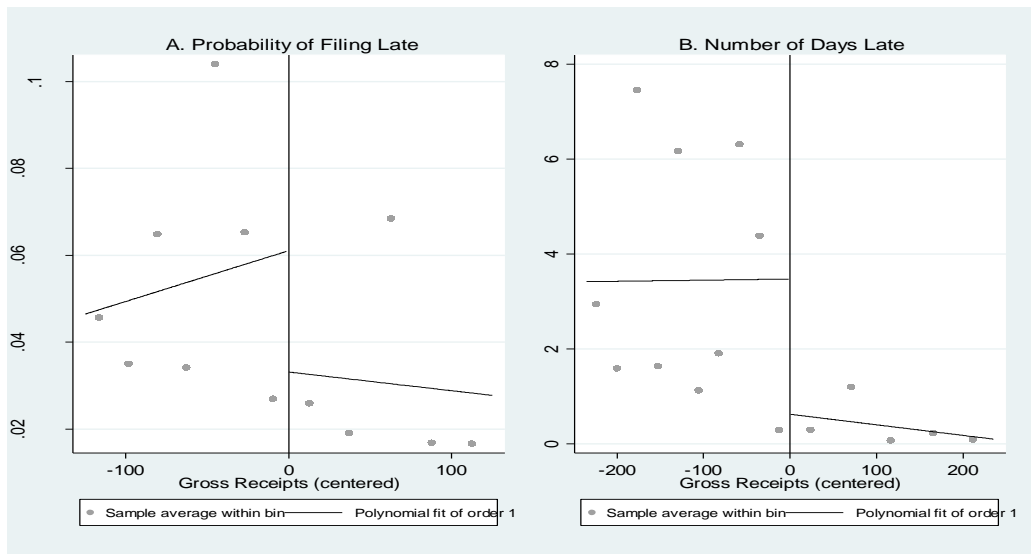
	Payment Compliance (Timeliness)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Probability of Paying Late						
LTO	-0.805 (0.891)	-0.704 (0.767)	-0.797 (0.966)	-0.657 (0.811)	-0.783 (0.810)	-0.723 (0.714)
Bandwidth (J\$ Millions)	290	290	270	270	310	310
Observations	874	797	793	720	972	890
Year Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes
Panel B - Number of Days Paid Late						
LTO	-77.79 (305.7)	-127.3 (282.1)	-90.51 (318.0)	-147.5 (292.0)	-64.15 (294.7)	-106.6 (273.1)
Bandwidth (J\$ Millions)	477	477	457	457	497	496
Observations	1,793	1,603	1,786	1,598	1,799	1,608
Year Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes
Panel C - Share of Reported CIT Paid on Time						
LTO	1.007 (0.870)	0.970 (0.795)	1.076 (0.971)	1.020 (0.883)	0.945 (0.780)	0.921 (0.724)
Bandwidth (J\$ Millions)	337	337	317	317	357	357
Observations	1,166	1,064	1,016	928	1,301	1,181
Year Dummy	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes
Panel D - Amount Paid on Time						
LTO	8.466 (13.22)	3.865 (5.190)	8.408 (13.61)	3.926 (5.302)	8.508 (12.91)	3.815 (5.105)
Bandwidth (J\$ Millions)	542	542	522	522	562	562
Observations	1,800	1,609	1,800	1,609	1,800	1,609
Year Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes

Note: This table presents estimates for payment compliance – timeliness of payments, for the CIT. The dependent variables in panels A-D are: the probability of paying late, the number of days late GCT is paid, the amount of GCT paid on time (in millions J\$) and the share of reported GCT paid on time. The results presented in columns 1-2 are for the optimal bandwidth proposed by Imbens and Kalyanaraman (2009), with and without controls respectively. Columns 3-4 present results for a smaller bandwidth (IK – J\$20 million) and columns 5-6 for a larger bandwidth (IK + J\$20 million). Controls include the ‘age’ of the taxpayer, the amount of estimated CIT paid, reported CIT and a dummy that capture whether or not the taxpayer operates in the financial sector. Regressions adopt a linear specification. Robust standard errors in parentheses.*p<0.10, **p<0.05, ***p<0.01

Filing Compliance for the GCT

Figure 5 presents RD plots of the effect of the LTO on filing compliance for the GCT. Panel A shows a discontinuous drop in the probability that taxpayers file late and panel B shows a drop in the number of days late that GCT returns are filed. Relative to the CIT, the figures show a more distinct separation in compliance behavior around the threshold and are more efficiently estimated.

Figure 5: Effect of the LTO on Timely Filing of the GCT



Note: Bandwidths are chosen to approximate the optimal (IK) bandwidth for the respective outcomes. RD Plots done using the default IMSE – evenly spaced (ES) method with spacing estimators Calonico et al. (2014b). The approach fits linear regression curves that approximate the conditional means of the outcome variable to the left and right of the cut off.

Table 11 presents estimates of the LATE for both filing compliance outcomes. The results in Panel A indicate a 22 percentage point reduction (column 1) in the probability of late filing. The results in Panel B suggest that taxpayers reduce the number of days late they file GCT returns by 14 days (column 1). Results are robust to the inclusion of controls and for larger and smaller bandwidth choices. In general the results suggest that taxpayers respond along both the

extensive margin – whether GCT returns are filed on time or not, and the intensive margin – how late GCT returns are filed.

Table 11: Effect of the LTO on Timely Filing of the GCT

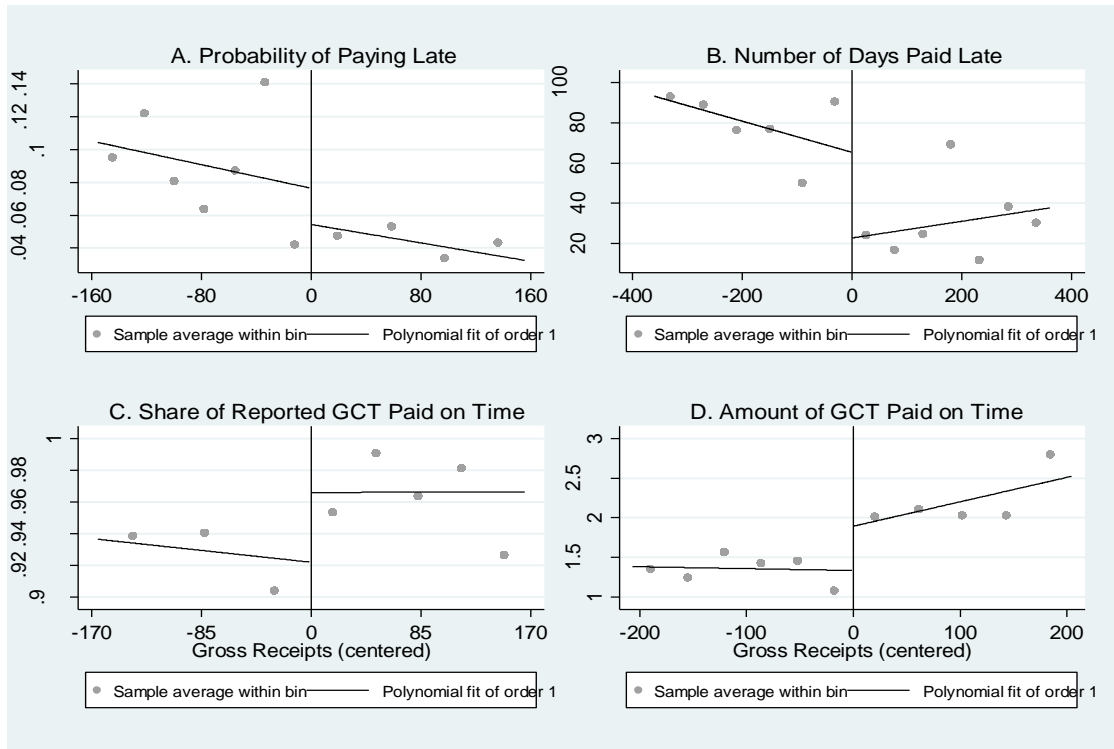
	Filing Compliance					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Probability of Filing Late						
LTO	-0.222** (0.0883)	-0.195** (0.0822)	-0.190** (0.0902)	-0.162* (0.0865)	-0.156*** (0.0482)	-0.150*** (0.0471)
Bandwidth (J\$ Millions)	125	125	105	105	145	145
Observations	6,026	5,849	4,934	4,786	13,345	12,977
Year-Month Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes
Panel B - Number of Days Late						
LTO	-14.44** (5.629)	-15.02*** (5.654)	-14.18** (5.823)	-14.53** (5.792)	-15.60*** (5.519)	-16.34*** (5.581)
Bandwidth (J\$ Millions)	238	238	218	218	258	258
Observations	12,790	12,433	11,190	10,879	14,385	13,985
Year-Month Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes

Notes: This table presents estimates for filing compliance for the GCT. The dependent variable in Panel A is an indicator equal 1 if the taxpayer filed late and 0 otherwise. The dependent variable in Panel B captures the number of days after the due date that taxpayers file a GCT return. The results presented in columns 1-2 are for the optimal bandwidth as proposed by Imbens and Kalyanaraman (2009), with and without controls respectively. Columns 3-4 present results for a smaller bandwidth (IK – J\$20 million) and columns 5-6 for a larger bandwidth (IK + J\$20 million). Controls include the ‘age’ of the taxpayer, amount of GCT arrears / credits and a dummy that capture whether or not the taxpayer operates in the financial sector. Regressions adopt a linear specification. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Payment Compliance for the GCT

Results for payment compliance for the GCT are graphed in Figure 6. Panels A and B show a discontinuous drop in the probability of paying late and the number of days late that taxpayers pay, respectively. Panels C and D show increases in the share of GCT and amount of GCT paid on time. We supplement these results with estimates from 2SLS regressions.

Figure 6: Effect of the LTO on Payment Compliance for the GCT



Note: Bandwidths are chosen to approximate the optimal (IK) bandwidth for the respective outcomes. RD Plots done using the default IMSE – evenly spaced (ES) method with spacing estimators Calonico et al. (2014b). The approach fits linear regression curves that approximate the conditional means of the outcome variable to the left and right of the cut off.

Table 12 gives the LATE for payment compliance for the GCT. Results in panel A (column 1) indicates a 17 percentage point reduction in the probability of paying late and panel B suggests a reduction in the number of days late by 247 days. Taxpayers’ ‘money response’ is captured in panels C and D, and show an increase in the amount of GCT paid on time by about J\$4 million and an increase in the share paid on time by roughly 27 percent. Results are robust to the inclusion of controls and alternative bandwidths. As in the case of filing, results for payment compliance suggest that taxpayers respond along both the extensive margin and intensive margins.

Table 12: Effect of the LTO on Payment Compliance for the GCT

	Payment Compliance (Timeliness)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Probability of Paying Late						
LTO	-0.174*	-0.150*	-0.173*	-0.141	-0.191**	-0.175**
	(0.0893)	(0.0826)	(0.0995)	(0.0928)	(0.0812)	(0.0751)
Bandwidth (J\$ Millions)	157	157	137	137	177	177
Observations	7,691	7,475	6,644	6,451	8,660	8,420
Year-Month Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes
Panel B - Number of Days Late						
LTO	-247.0***	-229.4***	-261.2***	-242.1***	-239.2***	-223.8***
	(49.23)	(48.79)	(53.07)	(52.01)	(45.89)	(46.01)
Bandwidth (J\$ Millions)						
Observations	25,693	25,040	22,631	22,036	29,547	28,809
Year-Month Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes
Panel C - Share of Reported GCT Paid on Time						
LTO	0.266**	0.193*	0.275*	0.200*	0.280***	0.216**
	(0.131)	(0.106)	(0.154)	(0.121)	(0.108)	(0.0993)
Bandwidth (J\$ Millions)	166	166	146	146	186	186
Observations	8,081	7,852	7,117	6,916	9,204	8,948
Year-Month Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes
Panel D - Amount of GCT Paid on Time						
LTO	4.415***	3.540***	5.044***	3.999***	3.994***	3.231***
	(0.830)	(0.696)	(0.997)	(0.810)	(0.740)	(0.634)
Bandwidth (J\$ Millions)	207	207	187	187	227	227
Observations	10,350	10,061	9,276	9,019	11,864	11,531
Year-Month Dummies	no	yes	no	yes	no	yes
Controls	no	yes	no	yes	no	yes

Notes: This table presents estimates for payment compliance for the GCT. The dependent variables in panels A-D are: the probability of paying late, the number of days late GCT is paid, the amount of GCT paid on time (in millions J\$) and the share of reported GCT paid on time. The results presented in columns 1-2 are for the optimal bandwidth as proposed by Imbens and Kalyanaraman (2009), with and without controls respectively. Columns 3-4 present results for a smaller bandwidth (IK – J\$20 million) and columns 5-6 for a larger bandwidth (IK + J\$20 million). Controls include the ‘age’ of the taxpayer, amount of GCT arrears / credits and a dummy that capture whether or not the taxpayer operates in the financial sector. Regressions adopt a linear specification. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

A contrast of the results for the CIT and GCT can offer insight into the potential for non-pecuniary factors more generally and taxpayer services is particular, to influence tax compliance in weak and stronger legal enforcement contexts. The null effects for the CIT and the positive compliance effects for the GCT provide suggestive evidence of a complementarity in the

relationship between the strength of the legal enforcement framework of the taxing regime and the provision of taxpayer services. To strengthen the case for comparability of compliance outcomes across the two tax types, we address possible confounding coming from heterogeneity across taxpayers in the CIT and GCT samples, by restricting the analysis to only taxpayers who file both CIT and GCT. We present results on filing and payment responses for this subsample in Table B.2 in Appendix B. The results are qualitatively similar to the baseline results using the optimal IK bandwidths – suggesting that a taxpayer responds more positively to the provision of services for the GCT – which has the stronger enforcement.

Comparability of the compliance outcomes may still be confounded by important differences in the structure of the taxes themselves. However we argue that the margins of response examined – timeliness of filing and payment conditional on reporting – are more comparably relative to other compliance outcomes typically studied such as reporting behavior. Still it can be argued, for example, that complexities of the CIT are more likely to cause delays in filing and payment compared to the GCT. But we show in panel B of Figure B.2 in Appendix B that even taxpayers who file and pay both taxes on time are still more compliant with the GCT – pay a larger share of reported tax liability. This reinforces our argument that the observed compliance behaviors across taxes are driven primarily by differences in the enforcement context, and mitigates concerns about confounding from sample selection and heterogeneity across taxes.

Heterogeneous Effects

Next we test for heterogeneous effects for the GCT. We focus on the GCT since the lax enforcement framework of the CIT will tend to nullify the potentially positive compliance effects of the LTO.⁴⁶ The compliance effects of the LTO can be very different across economic sectors for a number of reasons.⁴⁷ Differences in the business and regulatory environment across economic sectors that impose varied compliance requirements on firms can differentially affect taxpayer's compliance behavior. For example firms in the financial sector are already relatively heavily regulated to ensure they are 'fit and proper' and therefore may not be significantly impacted by the services provided through the LTO, compared to non-financial firms. A similar argument can be made for importing versus non-importing taxpayers, where the former require a tax compliance certification (TCC) in order to carry out its core business.⁴⁸ We examine differences in the compliance response across financial and non-financial sector taxpayers as well as importing and non-importing taxpayers.

Results for filing and payment compliance outcomes for financial and non-financial sector taxpayers are reported in Table B.3 of Appendix B. The results for financial sector taxpayers are reported in Panel A and indicate insignificant effects for all filing and payment compliance outcomes examined. The opposite is true for non-financial sector taxpayers, reported in Panel B, where we find significant improvements in all compliance outcomes examined. We find significant reductions in the probability of late filing (-0.17); the number of days late that GCT returns are filed (-15 days); the probability of paying late (-0.145); and the number of days

⁴⁶ Test for heterogeneity in compliance response for the CIT data yielded null results for all compliance outcomes.

⁴⁷ Studies that examine reporting behavior for the VAT for example point out that firms that sell directly to end users or consumers may be more likely to evade than firms who sell to other firms (Almunia & Rodriguez 2014; Pomeranz, 2013).

⁴⁸ We define importers as taxpayers for whom imports account for at least 20 percent of total supplies. The results are similar if we use higher qualification thresholds.

late that GCT is paid (-258 days). We also find significant gains in the amount of GCT paid on time (J\$2.4 million) and the share of reported GCT paid on time (16 percent). Comparing importing and non-importing taxpayers, we find a similar dynamic. The results presented in Table B.4 in Appendix B suggests that the compliance behavior of importers is not significantly impacted by the LTO, but for non-importers there are significant improvements in all the compliance outcomes examined. Estimates of the LATE for non-importers are qualitatively similar to that of non-financial taxpayers. These results suggest that compliance mechanisms – external to the tax administration, example using regulatory and other oversight bodies to ensure compliance or requiring TCC's to conduct certain business activities, are viable options to boost compliance.⁴⁹ In the context of this research, these external compliance mechanisms appear to act as substitutes for the provision of taxpayer services.

Conclusion

This essay examined the effect taxpayer service delivery through the LTO, on filing and payment of CIT and GCT for large taxpayers in Jamaica. We find generally positive compliance effects on filing and payment for the GCT but no effect for the CIT. A contrast of the results provides suggestive evidence of a complementarity in the relationship between the strength of the legal enforcement framework of the tax regime and the provision of taxpayer services. We argue that the relatively lax legal enforcement framework of the CIT moderates the potentially positive compliance effects of service provision. On the other hand, the stronger legal enforcement framework of the GCT complements the provision taxpayer services resulting in significant improvements in filing and payment. The null effects for the CIT raises doubts about the ability of tax morale factors and taxpayer services more specifically, by themselves, to

⁴⁹ The results presented in Tables 8 and 9 are from regressions using the optimal bandwidth (IK) only. The results are generally robust to adjustment in the bandwidth and the inclusion of controls.

significantly improve filing and payment compliance in developing countries. We also examined heterogeneous effects for key economic sectors and activities, and find improvements in filing and payment compliance for the GCT for taxpayers in the non-financial sector and for non-importers, but null effects for financial sector taxpayers and importers. Strict regulations of financial sector firms and TCC requirements for importers appear to be driving these results, and highlight the substitutability between external compliance mechanisms and the provision of taxpayer services.

The results have important implications for tax policy and tax administration in developing countries. The first points to limitations in the adoption of a softer approach to tax administration that attempts to encourage tax compliance by leveraging tax moral factors in a context of major enforcement deficiencies - example weak legal framework and corruption. Our results suggest that a strong(er) legal enforcement framework is required if non-pecuniary factors such as the provision of taxpayer services are to be effective in improving compliance. In the context of Jamaica, strengthening of the income tax law to remove or significantly limit opportunities for taxpayers to delay filing and payment of taxes could make the services provided by the LTO more valuable and could translate to improvements in compliance. Secondly, alternative compliance features, for example the use of strong regulatory or oversight bodies and requiring TCC's to conduct key business activities, are important drivers of tax compliance and can potentially serve as substitutes for other enforcement and non-pecuniary tax compliance programs carried on by the TAJ. Utilizing these 'compliance agents' could provide a low cost solution an already resource strapped tax administration in many developing countries.

Essay 3: The Effect of Public Goods Messaging on Personal Income Tax Compliance among the Self-Employed: Evidence from a Field Experiment in Jamaica

Introduction

Developing countries struggle to collect sufficient revenues to finance spending on important public services such as health, education, security, among others. This is evidenced by the relatively low tax-to-GDP ratios in these countries compared to their more advanced counterparts. The low tax effort in developing countries can be explained, partially, by low levels of tax compliance. Part of the problem stems from weak enforcement capacity, but other factors include perceptions of procedural fairness, existence of a non-compliance social norm and weak social and political institutions (Bird & Martinez-Vazquez, 2008; Besley & Persson, 2014). In light of this, to boost compliance, tax administrations are increasingly seeking to complement traditional enforcement approaches with those that leverage tax morale factors – such as social norm and public goods appeal. These complementary approaches are particularly important in developing countries where enforcement capacity is weak.

This essay implements a randomized messaging experiment to examine the effect of public goods messages on personal income tax (PIT) compliance among self-employed individuals in Jamaica. Traditionally, researchers use public goods messages to nudge behavioral nodes – aimed at eliciting tax compliance through reciprocity with respect to public service delivery. However, the effectiveness of these messages hinge on their ability to ‘re-couple’ in the mind of the taxpayer, the benefits of public goods and services provided by the government, with the amount of taxes paid to the government (Lamberton, De Neve, & Norton, 2014). Linked to the literature on charitable giving, is the view that knowledge - for example about the charity’s

activities and beneficiaries; and individual's ability to choose which charity to donate to, can increase generosity. Extending this thought to the tax compliance context, knowing how public spending is allocated and being able to influence spending allocation can build trust and provide taxpayer agency, ultimately leading to improvements in tax compliance (Djawadi & Fahr, 2013; Lamberton et.al; 2014).

We conduct two sub-experiments and examine the effect of public goods messages on two sets of compliance outcomes. In the first sub-experiment we examine the effect of the standard public goods message on payment of quarterly PIT obligations. Because there are no penalties for non-payment of quarterly PIT, the compliance effects will be free of confounding from enforcement action – a truly zero enforcement environment. In this context, the structure of the PIT penalty regime in Jamaica provides a unique opportunity to examine the effect of the standard public goods message on voluntary compliance. In the second sub-experiment we focus on payment of PIT arrears and expand the message context to include an augmented public goods message – which provides additional information on actual spending on key public goods and services. Providing information on spending allocation can 'open up the books', can elicit trust and arguably provide a sense of agency in the mind of taxpayers. In this sub-experiment the compliance outcomes relate to established PIT delinquencies, compared to sub-experiment one where there is no legal obligation to comply. Further, by augmenting the message content we broaden the scope of the analysis, in line with previous studies (Kirchler et al. 2008; Djawadi et al., 2013 Lamberton et al., 2014), to examine the incremental compliance effect of knowledge of government spending allocation.

We find that the standard public goods message had no effect on compliance with quarterly PIT payments in sub-experiment one. However results from sub-experiment two indicate positive effects, from both the standard and augmented messages, on the probability of making a payment and the amount of PIT arrears paid after nineteen weeks. However, point estimates from the standard and augmented public goods messages are not statistically different for any of the outcomes examined; suggesting that additional information on public spending allocations does not matter. Tests for heterogeneity in compliance response suggest that compliance gains are driven largely by older taxpayers. Unlike previous studies we find only limited evidence of differential compliance response between individuals with high and low tax debt.

Background

PIT applies to income above a tax exempt threshold. The PIT threshold was increased to J\$1,000,272 effective July 1, 2016, up from J\$592,800 in the previous year, making the effective threshold J\$796,536 for tax year 2016. For the period under study, PIT rates are 25 percent on income above the threshold, with a top marginal rate of 30 percent for income above J\$6 million.⁵⁰ Like most countries, the PIT in Jamaica can be characterized into two separate regimes. Most PIT revenue come from pay as you earn (PAYE) individuals and are withheld at source by employers. Individuals who earn only PAYE income are not required to file an income tax return and it is estimated that there were just fewer than 450,000 such individuals in the tax net in 2015. The second category of PIT revenue come from self-employed individuals, who are required to self-report earnings to the tax administration when they file their income tax each year. Based on data from the Statistical Institute of Jamaica (STATIN) the employed labor force

⁵⁰ The marginal PIT rate was increased to 30 percent for incomes at or above J\$6 million, effective July 1, 2016.

was estimated at roughly 1.1 million persons in January 2015, suggesting that there are approximately 650,000 self-employed individuals.

Jamaica's tax-to-GDP ratio averaged about 26 percent between 2000 and 2014. This is about 5 percentage points higher than the average for middle and upper middle income countries. The relatively creditable tax-to-GDP ratio may underplay the revenue mobilization challenges in Jamaica, compared to similarly situated economies. Notably, the critical issue has more to do with the distribution of the tax burden – who are called to bear the brunt? In a context of reasonably high tax ratios but where one segment of the tax population is made to bear the brunt of the burden, issues of economic inefficiencies and distributional inequities are even more pronounced. There is a long standing concern that the tax system places an inordinate burden on the PAYE employee. PAYE revenues were about J\$72 billion or 55 percent of total income and profit tax revenues in FY 2015/16. This compares to PIT collected from the self-employed of about J\$4 billion or 3.5 percent of total income and profit taxes.⁵¹ This imbalance is due at least in part to higher levels of tax evasion among the self-employed since they not only have greater opportunities to cheat (Kleven, Knudsen, Kreiner, Pedersen, & Saez, 2011) but perhaps also operate within a context where information or knowledge about how to cheat is more pervasive (Paetzold & Winner, 2014).⁵²

The self-employed are required to file estimated PIT returns during the first quarter of the current tax year, and make quarterly payments by March, 15, June, 15, September 15 and December 15. However, nonpayment of quarterly PIT does not attract a penalty and thus compliance is totally voluntary. In the absence of any formal punishment for non-compliance,

⁵¹ Data on PIT from the self-employed includes PIT collected from pensioners, in which case the burden carried by the self-employed may be even smaller.

⁵² PAYE taxpayers who earn other income are required to file income tax returns in respect of this income.

TAJ may need to leverage tax moral or non-pecuniary factors more generally – in an effort to boost compliance with quarterly payments. Annual returns and final PIT payments are due by March 15 of the following tax year. TAJ has long lamented the low levels of compliance among self-employed. Only 15 percent of registered self-employed taxpayers filed a tax return for tax year 2015, and only about half of these actually filed on time. With regards to payments, less than 1 percent of registered taxpayers paid any PIT in 2015. Principal PIT arrears were approximately 2.7 billion in April, 2016 or 57 percent of total projected PIT from the self-employed in fiscal year 2016/17.⁵³ Compared to quarterly payment obligations, the requirement to file and pay PIT, by the annual deadline, attracts a penalty.⁵⁴ Using non-pecuniary interventions could boost compliance where they are complementary to already established enforcement mechanisms.

Literature Review

The seminal work on tax compliance is due to Allingham and Sandmo (1972). Their model predicts that tax evasion rates fall as the probability of detection and the degree of punishment increases. Many authors have attempted to empirically validate the AS model, with varying success. A common finding is that the standard model predicts far too little compliance compared to that observed around the world (Torgler, 2004). However, the AS framework performs much better among individuals with self-reported income (Kleven, Knudsen, Kreiner , Pedersen, & Saez, 2011) where estimates of evasion come closer to the models predictions.

Complementary theories of tax compliance attempt to overcome some of the shortcomings of the AS approach. In particular, these theories posit that non-pecuniary or tax morale factors as well

⁵³ This figure includes amounts filed on returns but not paid, audit assessments and estimated assessments done by TAJ, but does not include interest and penalty balances. Total PIT arrears inclusive of interest a penalties are estimated at about J\$6 billion as at April 2016.

⁵⁴ Failure to file attract a penalty of J\$5,000 or up to 12 months in prison and nonpayment attracts interest of 40 percent annually.

as individuals subjective audit probabilities play an important role in explaining taxpayer behavior (Luttmer & Singhal, 2014)

Empirical work examining the determinants of tax compliance is still evolving. More recent studies that adopt quasi-experimental and experimental approaches in an attempt to uncover causal factors have emerged as the benchmark in the field. Blumenthal et al. (2001) find in the case of Minnesota that informing of the underlying tax compliance social norm as well as the range of public services financed by taxes has no effect on taxpayer's reported income or tax liability. Similarly, Torgler (2004, 2012) field experiment in Switzerland leveraged the importance of taxes to finance local public goods as well as the democratic freedom enjoyed by citizens, but found that such moral appeals had no impact on timely filing and payment of taxes, nor did it impact compliance as captured by reported income and claims for tax credits. These findings are in line with the narrative in much of the recent literature, which suggests that moral suasion may be less effective in encouraging compliance relative to enforcement factors.

However, (Bott, Cappelen, Sorensen, & Tungodden, 2014) find evidence in support of moral suasion – both social norm and public goods appeals, as an important determinant of compliance in Norway. Their experiment implements both an enforcement and moral appeal treatment, and finds that both had a significant and positive effect (of similar magnitude) on the reporting of foreign income. Interestingly they find that moral suasion impacted the intensive margin – people reported more foreign income, and the enforcement treatment impacted the extensive margin – more people reported having foreign income. A particularly interesting finding comes from Del Carpio (2013) who reports large and significant improvements in property tax compliance in Peru, from social norm messages, but null effects from enforcement messages. Interestingly she also finds evidence of possible crowd out of the pure morale effects

in treatments that interact both the social norm and enforcement messages – with estimated coefficient on the interacted messages being smaller than that on pure norm message. Notably, the effect of the social norm messages were largest during the period after increased municipal enforcement activity – arguably suggestive evidence of a complementarity in the compliance relationship between norm messaging and municipal enforcement action. More evidence on the positive compliance effects of moral suasion come from Hallsworth et.al. (2014). They find positive and significant compliance effects on payment of tax debts, from social norm and public goods messages in the UK. They advance that framing messages to increase moral cost can reduce procrastination and therefore increase tax payment – which in the case of the UK experiment resulted in an additional £3 million being collected within 23-days.

Recent lab experiments push beyond the standard public goods treatment and examine the effect of taxpayer knowledge about spending allocations and ability to allocate spending across key public goods and services, on tax compliance.⁵⁵ Djawadi and Fahr (2013) and Lambertson et al., (2014) find that providing information on public expenditures and allowing taxpayers to choose how their taxes are allocated significantly improves compliance. Both studies find larger compliance effects for the latter. Interestingly, Djawadi and Farh (2013) report that the positive compliance effects hold only when enforcement is weak – suggesting that these interventions may be substitutes for enforcement.

The contribution of this paper is three fold. Firstly we add to a nascent but growing literature that uses messaging experiments to study casual determinants of tax compliance in developing countries (Del Carpio, 2013; Castro & Scartascini, 2013; Ortega & Sanguinetti, 2013; Pomeranz, 2013; Carillo et al., 2014; Brockmeyer et al., 2016). This is the first in an

⁵⁵ Evidence from research on charitable giving suggest that persons are more willing to give if they know who the beneficiaries are (Cryder et al., 2013; Bohnet & Frey, 1999)

English speaking Latin American Caribbean (LAC) country. Secondly our enquiry into the potential effects of tax knowledge (as a proxy for taxpayer agency) on compliance outcomes moves what has so far been examined in the context of control lab environments, into the field. Thirdly the compliance outcomes we examine offers an opportunity to compare the effect of public goods messages (and non-pecuniary factors more generally) on compliance with a purely voluntary obligation – for which there are no penalties for non-compliance, and obligations for which non-compliance attract significant fines. Such a comparison might inform the nature of the relationship between the messages and the enforcement strength of the tax regimes.

Data and Experimental Design

Data

We use administrative data for over 100,000 registered self-employed individuals (as at July 18, 2016) in Jamaica.⁵⁶ From the universe of registered self-employed taxpayers we select our experimental sample of 17,073 ‘active’ taxpayers – defined as any person who either filed an annual income tax return for tax year 2015, an estimated return for the current tax year (2016) or who show up as having PIT arrears. The experimental sample for sub-experiment one comprise 3,368 active taxpayers with reported statutory income greater than the current effective tax exempt threshold of J\$796,536 for tax years 2015 and 2016, but with PIT arrears less than J\$20,000. The experimental sample for sub-experiment two comprise some 13,695 active taxpayers with PIT arrears greater than or equal to J\$20,000.⁵⁷

⁵⁶ All data are provided by TAJ. To maintain confidentiality, individual’s taxpayer registration numbers (TRN) were replaced by a system generated ‘customer key’ that is used to identify him or her throughout the experiment.

⁵⁷ In the context of this experiment, TAJ’s priority was to improve payments of outstanding PIT arrears. Consequently the sample for sub-experiment two was selected first and the remaining taxpayers who met the sample selection criteria for sub-experiment one were selected into that sub-experiment. Taxpayers were either selected into sub-experiment one or two, none were selected for both.

From the tax returns we have information on reported income, statutory income and amount of PIT filed for 2015 and 2016.⁵⁸ Data on PIT payments are from the payments ledger and cover all PIT payments made for 2016. Data on PIT arrears and a risk ranking for each taxpayer are from the compliance management dataset. Data on taxpayer characteristics such as age, sex, marital status and occupation are from the taxpayer registration number (TRN) database. Table 13 presents summary statistics for experimental samples from sub-experiments one and two.

Table 13: Summary Statistics Sub experiment One and Two

<i>Panel A: Sub-experiment One</i>			
Variables	Observations	Mean	Std. Dev.
IIT Paid Third Qrt (Dummy)	3,368	0.29	0.45
Amount of IIT Paid Third Qrt	3,368	54,641.44	237,476.30
IIT Paid Fourth Qrt (Dummy)	3,368	0.29	0.45
Amount of IIT Paid Fourth Qrt	3,368	55,890.92	240,320.30
Public Goods Message (stdPG)	3,368	0.724	0.447
Statutory Income	3,368	3,836,010.000	6,681,613.000
Risk	3,368	1.698	0.489
Age	3,368	51.269	13.831
Single	3,368	0.458	0.498
Male	3,368	0.607	0.489
Email	3,368	0.315	0.464
PIT Arrears	3,368	736.077	2,982.757
<i>Panel B: Sub-experiment Two</i>			
Variables	Observations	Mean	Std. Dev.
IIT Arrears Paid (Dummy)	13,695	0.08	0.28
Amount of PIT Arrears Paid	13,695	7,541.452	68,347.160
Public Goods Message (stdPG)	13,695	0.210	0.408
Public Goods & Spending Information (augr)	13,695	0.208	0.406
Risk	13,695	1.868	0.364
Age	13,695	48.180	12.087
Single	13,694	0.682	0.466
Male	13,695	0.730	0.444
Email	13,695	0.038	0.190
PIT Arrears	13,695	250,817.600	1,307,900.000

Experiment Design

We conduct two sub-experiments and examine the effect of public goods messages on two sets of compliance outcomes. Letters with salient messages were crafted in cooperation with the TAJ's communications team along with other members of TAJ's management team. All letters

⁵⁸ Data for 2016 are from estimated PIT returns filed. Final PIT returns for 2016 are due March 15, 2017.

were done on official TAJ letter head and sent in official Government of Jamaica envelopes using standard mail. Unlike previous studies that use registered mail, the use of standard mail is fitting to Jamaica's context as taxpayers are less likely to accept letters from the tax administration if sent through registered mail. This highlights a major challenge for TAJ, where effective communication, through letters is compromised because of low delivery rates.⁵⁹ Consequently, critical to this experiment is our ability to identify letters that were returned and not delivered to the taxpayer. To make it easy to identify returned mail relating to this experiment, envelopes were specially marked with a red outline of the Jamaican flag.⁶⁰ Table C.1 in Appendix C outlines the timeline for the experiment.

Sub-experiment One: Effect of Public Goods Messaging on Third and Fourth Quarter PIT Payments.

In sub-experiment one we examine the impact of the standard public goods message on individuals' payment of quarterly PIT obligations. Quarterly PIT payments are due on March, 15, June, 15, September 15 and December 15 of the current tax year. Letters were mailed between August 3 and 5, 2016 and thus were targeted at influencing payments for the third and fourth quarters.⁶¹ The salient message highlights the importance of taxes as a source of funding for important public services such as health care, education, security, among others.

⁵⁹ In recent years TAJ has rolled out an e-service platform geared toward facilitating 'real time' interaction with the taxpayer and provide convenient filing and payment options. TAJ has also increased its presence on social media, and has an active presence of Face Book and Twitter.

⁶⁰ Figure C.1 in Appendix C shows the envelope in which the letters were sent.

⁶¹ It is estimated to take between 3 – 5 working days for the letters to be delivered. Individuals are expected to receive these letters by the latest August 12, 2016 – one month before the September 15 deadline to make third quarter payments.

Standard Public Goods Message (stdnPG) - Treatment One (T1):

“Paying your taxes provide the funding necessary to pay for important public services such as health care, education, national security and other important services. If all self-employed individuals pay their income tax liabilities in full, this could provide funding to significantly increase spending in these important areas.”

The letter also reminded taxpayers to file and pay their income tax – both quarterly and annual obligations; the deadlines by which these actions must be taken; and a call to action or statement encouraging taxpayers to honor their obligation on time and in full. Information about where taxes can be filed and paid, and contact information – TAJ’s website and customer care help line in case there are questions or comments – were also included as standard text in all letters sent as a part of the experiment. Figure C.2 in Appendix C gives a sample of the letter used in this sub-experiment.

The sample for sub-experiment one comprised individuals who in either tax year 2015 or 2016 – based on estimated returns, reported statutory income greater than the effective annual tax exempt threshold of J\$796,536, but with PIT arrears less than J\$20,000.⁶² Individuals meeting the selection criteria (N = 3,368) were randomly selected into a treatment group (n = 2439) to receive a letter, and a control group (n = 929) that did not received a letter.⁶³

⁶² We use the effective tax exempt threshold for tax year 2016 as the selection criteria for the experimental sample.

⁶³ The larger sample size for the treatment group was chosen to compensate for the expected low delivery rate of letters. In the end however the delivery rate was much higher than expected. This might be on account of the fact that the taxpayers in this sub-experiment are particularly active and current.

Sub-experiment Two: Effect of Public Goods Messaging on Payment of Personal Income

Tax Arrears

In sub-experiment two, individuals are randomly assigned to one of two treatment groups that receive either the standard public goods message (T2) or an augmented public goods message (T3), or a control group that receives no letter.⁶⁴ These letters are targeted at individuals with tax arrears and were mailed between August 12th and 19th.⁶⁵ All letters notify the taxpayer that he or she is “... *among a group of taxpayers identified as having outstanding income tax liabilities*” and urge taxpayers to “... *pay all outstanding income tax liabilities*” if they have not already done so.⁶⁶ The salient message sent to those assigned to T2 is similar to T1 from sub-experiment one – but targets payment of tax arrears.

Standard Public Goods Message (stndPG) - Treatment Two (T2):

“Paying your taxes provide the funding necessary to pay for important public services such as health care, education, national security and other important services. If all self-employed individuals pay their income tax liabilities in full, this could provide funding to significantly increase spending in these important areas”.

Those assigned to T3 received a similar message, with the experimental variation coming from the inclusion of a sentence that provides information about the actual allocation of public

⁶⁴ Individuals selected for sub-experiment one were not included in this sub-experiment.

⁶⁵ The length of time taken to mail these letters was as a result of a ‘breakdown’ in the machine used to stamp and mark the envelopes in which the letters were to be sent. About half of the letters sent in the sub-experiment had been sent to the central post office for stamping before being mailed off which resulted in some delays.

⁶⁶ TAJ was careful in not stating how much the taxpayer owed since there may be questions about the accuracy of the figure due to system updating issues for example.

spending in the important areas of health, education and national security during the previous year.⁶⁷ The salient message sent to individuals in T3 reads:⁶⁸

Augmented Public Goods Message (augmPG) - Treatment Three (T3):

“Paying your taxes provide the funding necessary to pay for important public services such as health care, education, national security and other important services. Last year the Government spent approximately J\$54 billion on health, J\$90 billion on education and J\$57 billion on national security. If all self-employed individuals pay their income tax liabilities in full, this could provide funding to significantly increase spending in these important areas.”

The sample of individuals (n = 13,695) used in this sub-experiment are those with outstanding income tax liabilities on record of over J\$20,000 but less than J\$50 million.⁶⁹ Individuals meeting these criteria were randomly selected into the respective treatment and control groups. About 2,882 individuals were selected in treatment group (T2), 2,854 into treatment group (T3) and 7,959 were assigned to the control group.

Randomization Balance

To assess the validity of the experimental design we examine the balance of key taxpayer characteristics across the treatment and control groups for sub-experiments one and two. These characteristics include declared statutory income, amount of PIT arrears, risk rating (assigned by TAJ), whether the taxpayer has an email address on file, gender, marital status and age. Table 14

⁶⁷ The information on government spending in the various areas are factual and are published in the estimates of expenditure accompanying the budget for the current fiscal year 2016/2017.

⁶⁸ Figures C.3 and C.4 in the Appendix show the complete text for the letters sent to taxpayers in T2 and T3.

⁶⁹ The cutoff of J\$20,000 was chosen arbitrarily, but represents a non-trivial amount for the tax administration. We restrict the sample to tax arrears less than J\$50 million to mitigate potential bias from taxpayers located in the extreme tails of the distribution of tax arrears.

shows the mean covariate balance across the treatment and control group for sub- experiment one. Column 3 shows the difference between mean values of key taxpayer characteristics. The p values in column 4 suggests that relative to the control group, taxpayers in the treatment group are less likely to have an email address on file (p= 0.027), but are balanced across all other characteristics.

Table 14: Covariate Balance for Sub-experiment One

Covariates	Treatment (T1) (1)	Control (T1) (2)	T1 - C1 (3)	P- value (4)
Statutory Income (J\$ millions)	3.652	3.906	-0.254	0.326
Email	0.304	0.343	-0.039	0.027
Male	0.613	0.590	0.023	0.213
Single	0.500	0.474	0.026	0.248
Risk Rating	1.695	1.706	-0.011	0.568
Age	51.449	50.797	0.652	0.222
Observations	2,439	929		

Notes: Table shows the mean covariate balance for key taxpayer characteristics for sub-experiment one. Column 3 shows the difference in means for the treatment and control group and column 5 gives the associated p-values from the standard Wald test.

Covariate balance across the treatment and control groups for sub-experiment two is reported in Table 14. Columns 4 and 6 show the difference in means between the control group and T2 and T3, respectively. Relative to the control group, taxpayers in T2 are younger (p = 0.08) and taxpayers in T3 are more likely to be female (p = 0.05), but are balance across all other characteristics. Comparing the two treatments, taxpayers in T3 are younger relative to those in T2. We control for any imbalance in taxpayer characteristics by including them as covariates in our regression model. Notably however, the samples are balanced across all other characteristics.⁷⁰

⁷⁰ Kolmogorov-Smirnov test of equality of distribution for the continuous control variables reinforce the mean comparison in all cases except when comparing PIT arrears in T2 with the control group, in sub-experiment two.

Table 15: Covariate Balance for Sub-experiment Two

Covariates	Treatment (T2)	Treatment (T3)	Control (C2)	T2 - C2	p(value)	T3 - C2	p(value)	T2- T3	p(value)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Income Tax Arrears J\$	239,309.00	272,517.60	247,203.50	-7,894.50	0.77	25,314.10	0.40	-33,208.60	0.32
Statutory Income J\$	26,942.00	27,874.07	23,330.95	3,611.05	0.35	4,543.12	0.21	-932.07	0.57
Email	0.04	0.04	0.04	0.01	0.16	0.00	0.66	0.01	0.13
Male	0.74	0.71	0.73	0.00	0.75	-0.02	0.05	0.02	0.06
Single	0.67	0.68	0.69	-0.02	0.08	0.00	0.79	-0.02	0.22
Risk Rating	1.87	1.87	1.87	0.00	0.77	0.00	1.00	0.00	0.81
Age	48.20	48.05	48.22	-0.02	0.93	-0.17	0.52	0.15	0.64
Observations	2,882.00	2,854.00	7,959.00						

Notes: Table shows the mean covariate balance for key taxpayer characteristics for sub-experiment two. Columns 4 and 6 show the difference in means between the control group and the standard and augmented public goods treatment groups respectively. Columns 5 and 7 give the associated p-values from the standard Wald test.

Empirical Strategy

We examine payment compliance effects along an extensive and intensive margin, for both sub-experiments. We estimate the extensive margin treatment effects using the following probit model.

$$(Prob\ y_{i=1} | T; X_i) = \phi(\beta_0 + \beta_1 \sum_i T_i + X_i' \beta_2 + \epsilon_i) \quad (15)$$

For sub-experiment one, the dependent variable is a dichotomous variable set equal to one if the taxpayer made her third or fourth quarter payment by the stipulated deadline. For experiment two, the dependent variable equals one if the taxpayer paid any portion of his tax arrears by December 31 – roughly nineteen weeks after the estimated delivery date for the last batch of letters.

Next we estimate the treatment effects for the continuous outcomes. Our continuous outcome variable from sub-experiment one is the amount of PIT paid in the third and fourth quarter. In sub-experiment two we examine the effects on the amount of PIT arrears paid. Our baseline estimates are done using the following ordinary least squares (OLS).

$$Y_i = \beta_0 + \beta_1 \sum_i T_i + X_i' \beta_2 + \epsilon_i \quad (16)$$

The treatment indicators (T) correspond to the respective public goods messages sent to individuals in each sub-experiment and X_i is a vector of controls which include declared statutory income, PIT arrears, risk rating (assigned by TAJ), whether the taxpayer has an email address on file, gender, marital status and age.

Results

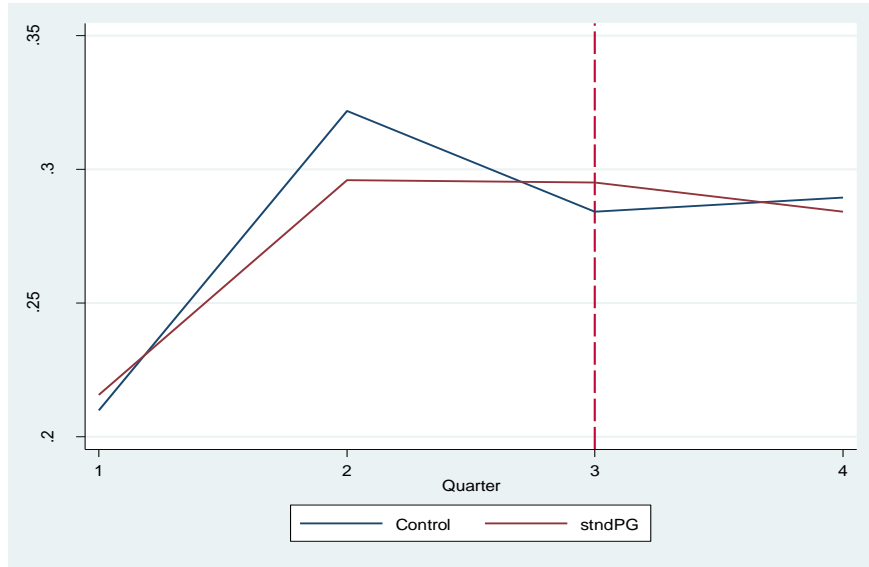
Sub-experiment One

We begin by estimating treatment effects for payment of third and fourth quarter obligations for tax year 2016. Figure 7 shows the payment rate for taxpayers in the treatment and control group for each quarter.⁷¹ Payment rates are low in the first quarter, increase sharply in the second and levels off in the third and fourth quarters. In general the payment rates for the treatment and control group are similar pre and post treatment. – suggesting the treatment had no effect. Figure 8 shows average quarterly PIT payments for the treatment and control group before and after treatment. Average quarterly PIT payments for the treatment group are always larger than the control, with the largest differences showing up in the second and third quarters, with some convergence by the fourth quarter.⁷²

⁷¹ The payment data does not link PIT paid to a specific quarter, but rather payments are linked to the year of assessment. The data however does provide a payment date and we assume that payments received within a particular quarter are intended for that quarter's liabilities.

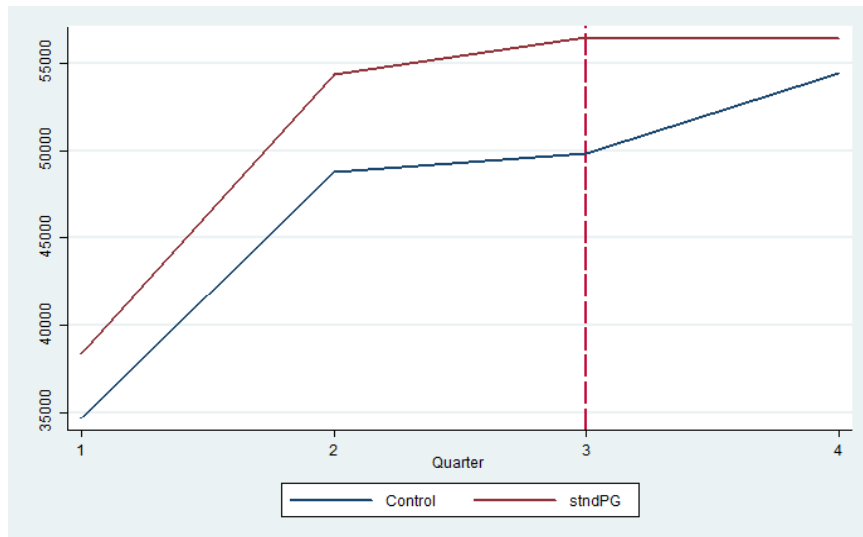
⁷² The pre-treatment difference between the treatment and control groups are not statistically significant.

Figure 7: PIT Payment Rates by Quarter



Notes: Figure shows PIT payment rates for each quarter in tax year 2016. Quarterly payments are due March 15, June 15, September 15 and December 15 of the tax year. The payment rate captures the proportion of sampled taxpayers in the treatment and control groups that made a payment during each quarter. Payments data are not specific to the quarter and so in calculating the payment rate we assume that payments made within a particular quarter are intended for that quarter.

Figure 8: Average PIT Payment by Quarter



Notes: Figure shows average PIT payments (J\$) for the treatment and control group for each quarter in tax year 2016. Quarterly payments are due March 15, June 15, September 15 and December 15 of the tax year. Payments data are not specific to the quarter and so we assume that payments received within a particular quarter are intended for that quarter.

Table 16 reports estimates for the probability of third (columns 1-4) and fourth (columns 5-8) quarter payments. Results from both probit and linear probability models indicate positive but statistically insignificant effects in the third quarter and negative but insignificant effects in the fourth quarter. Results are robust to the inclusion of controls and alternative specifications of the underlying regression.

Table 16: Probability of Making Quarterly PIT Payments

	Probability of Payment - Third Quarter				Probability of Payment - Fourth Quarter			
	LPM		Probit		LPM		Probit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Public Goods	0.0110 (0.0174)	0.0149 (0.0170)	.0110707 (0.0176)	0.0162 (0.0177)	-0.00543 (0.0175)	-0.00111 (0.0172)	-0.0054 (0.0174)	-0.0004 (0.0175)
Risk Rating		-0.127*** (0.0163)		-0.1285*** (0.0160)		-0.0716*** (0.0162)		-0.0722*** (0.0159)
Age		0.00134** (0.000665)		0.0014** (0.0007)		0.000238 (0.000677)		0.00031 (0.0007)
Single		-0.0266 (0.0182)		-0.0283 (0.0190)		-0.0213 (0.0184)		-0.0218 (0.0190)
Male		0.0279* (0.0158)		0.0293* (0.0165)		0.0248 (0.0160)		0.0256 (0.0164)
Email		0.187*** (0.0174)		0.1842*** (0.0167)		0.159*** (0.0174)		0.1550*** (0.0165)
Constant	0.284*** (0.0148)	0.365*** (0.0507)			0.290*** (0.0149)	0.340*** (0.0508)		
Observations	3,368	3,368	3,368	3,368	3,368	3,368	3,368	3,368

Notes: Columns 1-4 in the table present estimates of the probability of making third quarter payments (September, 15) from linear probability and probit models and columns 5-8 present estimates for the fourth (December, 15) quarter. Coefficient estimates from probit regressions give the marginal effects. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table 17 present results for the amount of PIT paid in each quarter. We present results from OLS regressions of the treatment on the outcome variable – in levels, as well for an appropriate transformation of the outcome variable. Because the outcome variable, the amount of PIT paid, is ‘zero-inflated’, we present estimates of the treatment effects using an inverse hyperbolic sign (IHS) transformation. The IHS is an alternative to the log transformation, with

the advantage that it is well defined over zero.⁷³ Columns 1-4 and 5-8 give results for PIT payments for the third and fourth quarters respectively. Similar to the extensive margin effects above, the treatment effects are statistically insignificant in both specifications for both quarters.

Table 17: Amount of Quarterly PIT Paid

	Third Quarter				Fourth Quarter			
	Amount Paid		IHS Trans		Amount Paid		IHS Trans	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Public Goods (stndPG)	6,662 (7,493)	5,916 (7,334)	0.102 (0.210)	0.144 (0.205)	1,989 (8,062)	1,036 (7,917)	-0.0842 (0.210)	-0.0382 (0.206)
Risk Rating		-30,734*** (10,357)		-1.536*** (0.199)		-22,827** (11,384)		-0.922*** (0.199)
Age		819.5*** (263.7)		0.0181** (0.00795)		446.8* (252.6)		0.00613 (0.00808)
Single		-4,817 (8,356)		-0.335 (0.219)		-12,455 (8,140)		-0.288 (0.221)
Male		17,533** (8,199)		0.409** (0.188)		19,721** (8,240)		0.373* (0.190)
Email		17,173** (7,354)		2.204*** (0.209)		9,253 (7,416)		1.885*** (0.209)
PIT Arrears		-2.281*** (0.419)		-1.71e-06 (2.86e-05)		-2.280*** (0.458)		3.01e-05 (2.98e-05)
Constant	49,817*** (5,322)	48,378* (25,875)	3.399*** (0.179)	4.263*** (0.607)	54,450*** (6,145)	63,499** (26,431)	3.446*** (0.179)	3.955*** (0.609)
Observations	3,368	3,368	3,368	3,368	3,368	3,368	3,368	3,368

Notes: Columns 1-4 and 5-6 in the table present estimates for the amounts of PIT arrears paid and its inverse hyperbolic sine (IHS) transformation, for the third and fourth quarters respectively. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01

Sub-experiment Two

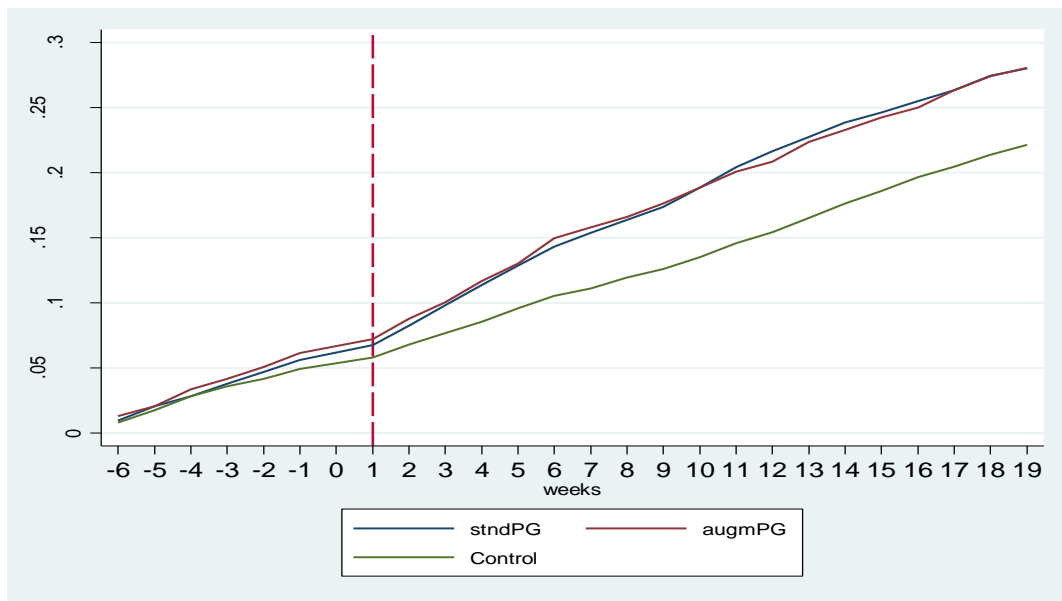
Figure 9 graphs the cumulative payment rate for PIT arrears, six weeks pre-treatment and nineteen weeks post treatment.⁷⁴ Pre-treatment payment rates are roughly the same for the treatment and control groups. The figure shows a small divergence between the control and treatment groups one week prior to the estimated delivery date for the last batch of letters. This is

⁷³ The inverse hyperbolic sine transformation (IHS) takes the form $\sin^{-1}(y) = \log(y + (y^2 + 1)^{1/2})$. See Burbidge et. al. (1988) for a more detailed discussion on the IHS transformation.

⁷⁴ The estimated delivery date for the last batch of letters for sub experiment two is August 26. The pre-treatment period runs from the date that the initial data were pulled (July 18) to the date of delivery of the last batch of letters (August 26). The post treatment period runs from the estimated delivery date to December 31, when payments data were provided.

likely due to treatment effects from letters sent in the first batch which would have been delivered to some taxpayers one week prior to the estimated delivery of the last batch of letters. In the week immediately after delivery of the last batch of letters, there is a noticeable increase in the payment rate for both treatment groups, relative to the control group. This difference increases progressively from the second week onward – suggesting material treatment effects. Notably however, payment rates for the standard and augmented public goods treatments are not statistically different, suggesting non-differential effects from having knowledge of public spending allocations.

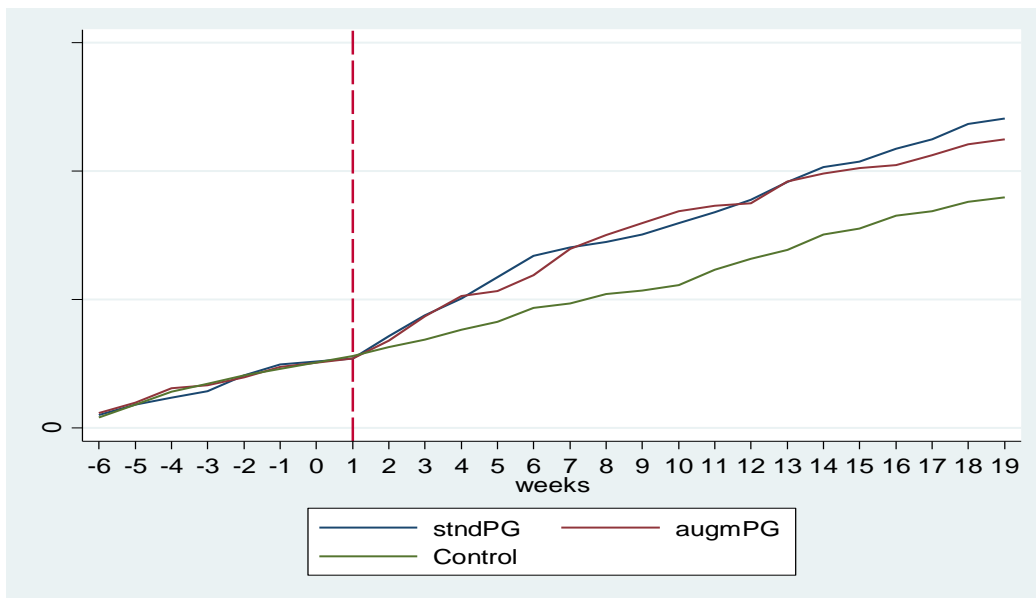
Figure 9: Payment Rate for PIT Arrears



Notes: Figure shows the cumulative payment rate for PIT arrears, 6 weeks before and 19 weeks after the estimated delivery date for the last batch of letters. The shorter pre-treatment period captures the time frame between the date of data extraction by TAJ and the estimated date of delivery of the last batch of letters. Payment rates are calculated as the cumulative share of persons in the respective groups that made a payment by each of the weeks highlighted. The red dashed line indicates the estimated date of delivery for the last batch of letters – experiment start date.

Figure 10 plots the cumulative average payments of PIT arrears for the treatment and control groups. Cumulative average PIT payments for the treatment and control groups are roughly the same for each of the six weeks prior to the start of the experiment— indicated by the dashed red line. Relative to the control group, cumulative payments for both treatment groups begin to pick up in the first week post-treatment. The gap between the control and treatment groups continues to widen through week ten before stabilizing. In general the non-parametric estimates suggest that both public goods messages resulted in an increase in the average amount of PIT paid relative to the control group, and that the increase took place gradually over time, then slowed during the latter weeks. As in the analysis of payment rates above, the treatment effects for the standard and augmented public goods messages are not statistically different.

Figure 10: Cumulative Payment for PIT Arrears per Person



Notes: The figure shows cumulative average payment of PIT arrears per person for the treatment and control groups 6 weeks prior and 19 weeks after the estimated delivery date for the last batch of letters. The red dashed line indicates the estimated date of delivery for the last batch of letters – the experiment start date. Relative to the control group, average cumulative payments post treatment are higher for both treatment groups. However average cumulative payments rates do not differ across the two treatments.

We supplement the graphical analysis with more robust estimates from equations (15) and (16). We examine taxpayer’s response at roughly nineteen weeks post treatment. Table 18 present results for the probability of payment of PIT arrears – an extensive margin response. LPM results (column 1) suggest that the standard and augmented public goods messages increased the probability of payment by 2.7 and 2.4 percentage points respectively. Probit results are similar, and indicate a 2.6 and 2.3 (column 3) percentage point increase in payment from the standard and augmented public goods messages respectively. Results are robust to the inclusion of controls.

Table 18: Effect of Public Goods Messages on Probability of Payment of PIT Arrears

	Probability of Payment			
	LPM		Probit	
	(1)	(2)	(3)	(4)
Public Goods	0.0265*** (0.00630)	0.0249*** (0.00622)	0.0258*** (0.0058)	0.0237*** (0.0057)
Public Goods & Spending Information	0.0236*** (0.00626)	0.0238*** (0.00618)	0.0233*** (0.0059)	0.0230*** (0.0058)
Risk		0.0195*** (0.00661)		0.0186*** (0.0068)
Age		-0.000210 (0.000217)		-0.0003 (0.0002)
Single		-0.0206*** (0.00594)		-0.0200*** (0.0056)
Male		-0.00269 (0.00547)		-0.0023 (0.0053)
Email		0.202*** (0.0199)		0.1229*** (0.0091)
PIT Arrears		-6.68e-10 (1.42e-09)		-7.15e-10 (1.32e-09)
Constant	0.0731*** (0.00292)	0.0556*** (0.0188)		
Observations	13,695	13,694	13,695	13,694
Wald (T1 = T2) p-value	0.7142	0.8907	0.7142	0.9151

Notes: Table presents results for the effect of the standard and augmented public goods message on the probability of payment of PIT arrears after 19 weeks. Columns 1-2 present results from LPM estimates and columns 3-4 present results from probit regressions (marginal effects). Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table 19 summarizes results for the amount of PIT arrears paid – an intensive margin response. Results in column 1 indicate that the standard public good message increase the amount of PIT arrears paid by J\$2, 590 (US\$20). Point estimates for the augmented public goods message are qualitatively similar but are not statistically significant. Results are robust to the inclusion of controls (column 2). Similar to sub-experiment one we present alternative estimates of the treatment effects using the IHS transformation of the outcome variable. Results in column 3-4 indicate that relative to the control group, the standard and augmented public goods messages increase payment of PIT arrears by 29 and 25 percent respectively.

Table 19: Effect of Public Goods Messages on Payment of PIT Arrears

	Payment of PIT Arrears			
	Amount Paid		IHS Transformation	
	(1)	(2)	(3)	(4)
Public Goods (stndPG)	2,836*	2,590*	0.290***	0.271***
	(1,526)	(1,515)	(0.0699)	(0.0688)
Public Goods & Spending Info (augmPG)	2,267	2,166	0.245***	0.247***
	(1,734)	(1,720)	(0.0690)	(0.0682)
Risk		4,751***		0.253***
		(1,614)		(0.0738)
Age		26.32		-0.00162
		(64.19)		(0.00241)
Single		-6,203***		-0.257***
		(1,564)		(0.0663)
Male		-2,173		-0.0318
		(1,490)		(0.0609)
Email		28,045***		2.343***
		(5,732)		(0.232)
PIT Arrears		0.00384***		7.62e-09
		(0.00139)		(1.93e-08)
Constant	6,472***	194.3	0.803***	0.521**
	(654.9)	(4,807)	(0.0324)	(0.210)
Observations	13,695	13,694	13,695	13,694
Wald (stndPG = augmPG) p-value	0.7879	0.8396	0.6043	0.7784

Notes: Table presents results for the effect of the standard and augmented public goods messages on payment of PIT arrears after 19 weeks. Columns 1-2 present results for the amount of PIT arrears paid and columns 3-4 present results for the inverse hyperbolic sine (IHS) transformation of the amounts paid. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Relative to the standard public goods message, including information about public spending allocation does not have a differential effect on compliance. This is true for all outcomes examined and holds across different estimation strategies. This might suggest that self-employed individuals in Jamaica are less sensitive to public spending allocation decisions, but rather may be content with government's financing of general public services. Another potential explanation for this result is that the treatment may not have been strong enough to isolate the significance of spending in any one area, thus mitigating potentially positive compliance effects from increased transparency or taxpayer agency.

Effect of Treatment on the Treated (TOT) for Sub-experiment One and Two

The above estimates give the intent to treat (ITT) effect of the messages. These are arguably lower bound since they estimate compliance effects for taxpayers selected to receive a particular treatment without regard for whether they were actually treated. Because letters sent by the tax authority have a less than perfect delivery rate, we refine the baseline estimates above by accounting for undelivered mailings. We estimate the effect of treatment on the treated (TOT) using a two stage least squares (2SLS) approach. In the first stage we estimate the likelihood of actually receiving a letter conditional on being selected for treatment, and use the fitted values from the first stage to estimate the TOT effect in the second stage.

Tables 20 and 21 present TOT estimates for sub-experiments one and two respectively. As in the baseline, all the outcomes examined in sub-experiment one are statistically insignificant. Results for sub-experiment two indicate that relative to the control group the standard and augmented public goods message increase the probability of payment by 3.5 and 3.2 percentage points respectively (column 1). The standard public goods message increase the

amount of PIT arrears paid by J\$3,781 (US\$29) (column 3), but again the effect of the augmented message though positive is statistically insignificant. Results from the IHS transformation indicate increases of about 39 and 33 percent for the standard and augmented public goods messages respectively (column 5). The TOT estimates are slightly larger than the ITT estimates above and arguably more accurately capture the treatment effects.

Table 20: TOT Effect of Public Goods Message on Quarterly PIT Payment

	Third Quarter			Fourth Quarter		
	Payment Prob (1)	Amount Paid (2)	IHS Trans (3)	Payment Prob (4)	Amount Paid (5)	IHS Trans (6)
Public Goods (stdPG)	0.0167 (0.0189)	6,593 (8,163)	0.160 (0.228)	-0.00117 (0.0191)	1,155 (8,812)	0.160 (0.228)
Risk Rating	-0.128*** (0.0163)	-30,743*** (10,346)	-1.536*** (0.198)	-0.0750*** (0.0163)	-22,829** (11,373)	-1.536*** (0.198)
Age	0.00130* (0.000665)	804.5*** (261.8)	0.0177** (0.00795)	0.000249 (0.000677)	444.1* (255.4)	0.0177** (0.00795)
Single	-0.0267 (0.0182)	-4,804 (8,344)	-0.335 (0.219)	-0.0216 (0.0184)	-12,452 (8,128)	-0.335 (0.219)
Male	0.0278* (0.0158)	17,484** (8,191)	0.408** (0.188)	0.0250 (0.0159)	19,713** (8,222)	0.408** (0.188)
Email	0.187*** (0.0174)	17,185** (7,343)	2.204*** (0.208)	0.160*** (0.0174)	9,255 (7,404)	2.204*** (0.208)
PIT Arrears	1.34e-06 (2.53e-06)	-2.266*** (0.417)	-1.32e-06 (2.85e-05)	4.18e-06 (2.65e-06)	-2.277*** (0.456)	-1.32e-06 (2.85e-05)
Constant	0.368*** (0.0501)	49,161* (25,899)	4.282*** (0.599)	0.342*** (0.0502)	63,636** (26,434)	4.282*** (0.599)
Observations	3,368	3,368	3,368	3,368	3,368	3,368

Notes: Table summarizes treatment on the treated (TOT) estimates from 2SLS IV regressions, for the full model which includes controls. Columns 1-3 report results for the third quarter and gives results for the probability of payment (LPM), the amount paid and the IHS of the amounts paid. Columns 4-6 report results for the same outcomes for the fourth quarter. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table 21: TOT Effect of Public Goods Message on Payment of PIT Arrears

	Amount of IIT Arrears Paid					
	Prob of Payment		Amount Paid		IHS Transformation	
	(1)	(2)	(3)	(4)	(5)	(6)
Public Goods (stdnPG)	0.0353*** (0.00838)	0.0332*** (0.00828)	3,781* (2,034)	3,454* (2,019)	0.386*** (0.0930)	0.362*** (0.0916)
Public Goods & Spending Info (augmPG)	0.0318*** (0.00843)	0.0321*** (0.00833)	3,059 (2,339)	2,923 (2,320)	0.330*** (0.0930)	0.334*** (0.0919)
Risk		0.0204*** (0.00660)		4,836*** (1,609)		0.262*** (0.0737)
Age		-0.000250 (0.000216)		22.50 (64.50)		-0.00204 (0.00240)
Single		-0.0202*** (0.00593)		-6,158*** (1,566)		-0.252*** (0.0662)
Male		-0.00220 (0.00547)		-2,127 (1,499)		-0.0267 (0.0609)
Email		0.201*** (0.0198)		27,904*** (5,741)		2.327*** (0.231)
PIT Arrears		-8.69e-10 (1.41e-09)		0.00382*** (0.00138)		5.49e-09 (1.92e-08)
Constant	0.0731*** (0.00292)	0.0553*** (0.0188)	6,472*** (654.8)	165.6 (4,804)	0.803*** (0.0324)	0.518** (0.209)
Observations	13,695	13,694	13,695	13,694	13,695	13,694
Wald (T1 = T2) p-value	0.7430	0.9203	0.7993	0.8502	0.6301	0.8060

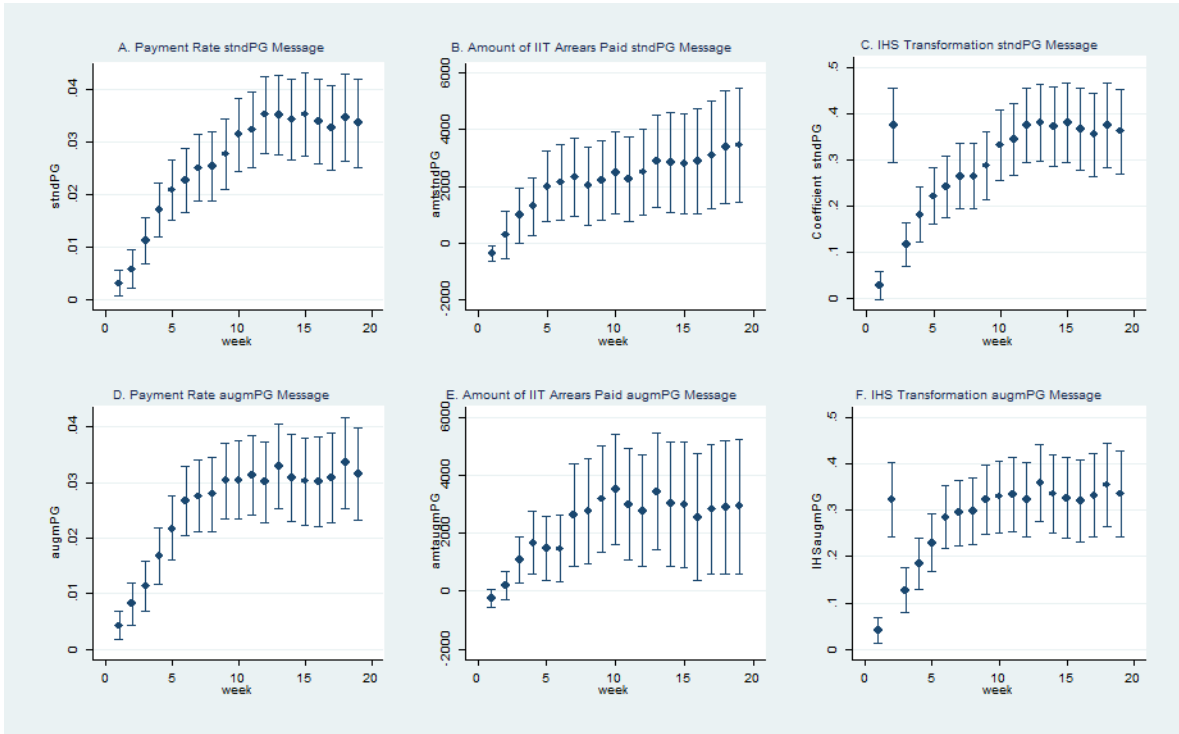
Notes: Table summarizes treatment on the treated (TOT) estimates from 2SLS IV regressions. Columns 1-2 report results on the probability of payment from linear probability models (LPM), columns 3-4 present results for the amount of PIT arrears paid in levels and columns 5-6 present results for the IHS transformation of the amounts paid. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Robustness and Heterogeneity

In addition to conditioning on taxpayer characteristics we conduct two additional sensitivity tests for sub-experiment two.⁷⁵ We do this by examining treatment effects overtime and by varying the ‘start date’ of the experiment and running a placebo test. Whether treatment effects are contemporaneous, whether they take hold over time or whether they dissipate over time is important for the tax administration. We examine variation in ‘short term’ treatment effects for the standard and augmented public good messages from sub-experiment two in Figure 11.

⁷⁵ Sensitivity tests for sub-experiment one confirm the null results from the baseline.

Figure 11: Coefficient Estimates of Treatment Effects over Time



Notes: Figures plot point estimates of treatment effects (TOT), with confidence intervals. The outcome variables are the payment rate, amount of PIT arrears paid and an IHS transformation of the amount paid. The outcome variables capture the cumulative payments by week from, week 1 to 19. Figures A – C show TOT effects from the standard public goods message (stdPG) and Figures D – F show TOT effects from the augmented public goods message (augmPG) for the same set of outcomes. All, except the coefficient on the amount of PIT arrears paid from the augmPG message are statistically significant starting from week 4 onwards.

Payment rates increase more sharply over the first 6 to 10 weeks post treatment and then stabilize over the ensuing weeks. The amount of PIT arrears paid follows a similar, but less distinct pattern. Point estimates from the IHS transformation show similar patterns, except we see a spike in the second week post treatment. This may be explained by a huge inflow of payments from both treatment groups relative to the control group by the second week post treatment, which was never replicated during the ensuing weeks. In general the sign and trajectory of the point estimates for the outcomes we examine are consistent with robust treatment effects over the post treatment period.

As an additional robustness test we examine treatment effects based on a placebo experiment start date, 6 weeks before the actual start date of the experiment, and examine compliance outcomes by the actual start date.⁷⁶ We compare these results with estimates of treatment effects 6 weeks after the actual start date. If the compliance outcomes identified above are due to the actual treatments, then the results from the placebo test should show null effects. The results in Table 22 confirms this, showing null treatment effects for all outcomes examined 6 weeks prior to the actual start date of the experiment, but positive and statistically significant effects for the payment rate and the IHS transformation 6 weeks after the actual start date, further validating our baseline results.

Table 22: Placebo Test – Payment of PIT Arrears Six Weeks Before and After Experiment

	Six Weeks Prior to Treatment			Six Weeks Post Treatment		
	Payment Prob (1)	Amount (2)	IHS (3)	Payment Prob (4)	Amount (5)	IHS (6)
Public Goods	0.00681 (0.00513)	-1,893 (2,175)	0.0699 (0.0556)	0.0225*** (0.00607)	2,142 (1,353)	0.241*** (0.0657)
Public Goods & Spending Information	0.00489 (0.00512)	-1,868 (2,211)	0.0531 (0.0559)	0.0275*** (0.00628)	1,471 (1,149)	0.284*** (0.0676)
Risk	0.0147*** (0.00394)	2,333*** (695.8)	0.175*** (0.0420)	0.00930* (0.00492)	1,429 (1,007)	0.121** (0.0540)
Age	3.49e-05 (0.000138)	-49.76 (63.62)	0.000551 (0.00152)	2.55e-05 (0.000163)	19.14 (35.18)	0.000695 (0.00176)
Single	-0.00671* (0.00378)	-1,013** (500.4)	-0.0862** (0.0411)	-0.00588 (0.00427)	-3,398*** (814.7)	-0.0883* (0.0468)
Male	0.00250 (0.00337)	1,408 (1,142)	0.0297 (0.0368)	-0.00446 (0.00405)	-608.8 (911.3)	-0.0501 (0.0441)
Email	0.110*** (0.0152)	7,755*** (2,055)	1.235*** (0.172)	0.121*** (0.0162)	11,442*** (3,539)	1.384*** (0.185)
PIT Arrears	1.57e-09 (1.41e-09)	0.00115** (0.000509)	2.54e-08 (1.72e-08)	9.12e-10 (1.22e-09)	0.00188** (0.000731)	2.23e-08 (1.68e-08)
Constant	-0.00159 (0.0112)	529.1 (2,781)	-0.0529 (0.120)	0.0180 (0.0142)	678.0 (2,657)	0.146 (0.155)
Observations	13,694	13,694	13,694	13,694	13,694	13,694
Wald (T1 = T2) p-value	0.7608	0.9742	0.8062	0.5228	0.6817	0.6128

Notes: Table presents estimates of compliance effects six (6) weeks before and after the experimental start date, for comparison purposes. Columns 1-3 show null effects for all outcomes in the pre -treatment period but positive and statically significant effects on payment probability and the IHS transformed variable in the post treatment period. *p<0.10, **p<0.05, ***p<0.01.

⁷⁶ We do not have data on PIT arrears prior to the July 18, 2016 – when the initial data were provided, and so are unable to run the placebo test for a longer pre-treatment period.

If messages impact taxpayers differentially, and given that such communication strategies can be costly to the tax administration, efficiency dictates that communications target taxpayers who are most likely to respond. We examine heterogeneity in treatment effect by age and level of tax arrears. There is some evidence that younger taxpayers are less compliant (Andreoni, Erard & Feinstein, 1998). Several hypotheses have been proffered to explain the relative non-compliance of youngsters, and include lack of tax knowledge, perception that paying taxes offer no benefit, more risk loving and rebellious.⁷⁷ The tendency for older taxpayers to be more compliant is explained by an “ageing out effect”. Larger tax debtors are also believed to be less likely to comply – since larger debt stock implies a penchant for non-compliant behavior.

Table 23 presents results for taxpayers below and above 35 years old.⁷⁸ Younger taxpayers did not respond to either treatment (columns 1-3) and that the positive compliance effects are driven by older taxpayers. The standard and augmented public goods messages increase the probability of payment by 3.9 and 3.2 percentage points (column 4), and the amount of PIT arrears paid by 43 and 34 percent (column 6) respectively. Only the standard public goods message had a significant effect, increasing the amount of PIT arrears paid, by J\$3, 951 (US\$31) (column 5).

⁷⁷ See Braithwaite, Smart and Reinhardt (2006) for a discussion on the age related determinants of tax compliance.

⁷⁸ The age cut off captures the population cohort labeled as ‘millennials’. In general the results are similar if we use the median age as the cut-off.

Table 23: Impact of Public Goods Messages on Payment of PIT Arrears by Age Cohort

	Age <= 35 years old			Age > 35 years old		
	Prob of Paying (1)	Amount Paid (2)	IHS Transformation (3)	Prob of Paying (4)	Amount Paid (5)	IHS Transformation (6)
Public Goods (stdPG)	-0.00495 (0.0220)	264.5 (1,912)	-0.0565 (0.239)	0.0390*** (0.00893)	3,951* (2,314)	0.426*** (0.0991)
Public Goods & Spending Info (augmPG)	0.0274 (0.0245)	12,554 (10,250)	0.267 (0.265)	0.0323*** (0.00882)	1,293 (2,102)	0.338*** (0.0977)
Risk	-0.00673 (0.0227)	4,445 (3,273)	-0.0604 (0.254)	0.0227*** (0.00688)	4,790*** (1,763)	0.290*** (0.0768)
Age	-0.00120 (0.00221)	-913.4 (1,303)	-0.00810 (0.0241)	-0.000192 (0.000257)	50.14 (71.80)	-0.00122 (0.00288)
Single	0.00738 (0.0298)	-131.4 (3,207)	0.0615 (0.332)	-0.0208*** (0.00607)	-6,243*** (1,624)	-0.260*** (0.0677)
Male	-0.0131 (0.0152)	-235.0 (2,933)	-0.156 (0.166)	-0.00118 (0.00589)	-2,635 (1,681)	-0.0139 (0.0658)
Email	0.105** (0.0524)	14,369* (8,393)	1.262** (0.605)	0.212*** (0.0213)	29,571*** (6,389)	2.450*** (0.249)
PIT Arrears	6.72e-09 (1.03e-08)	0.00473 (0.00365)	1.15e-07 (1.38e-07)	-1.24e-09 (1.42e-09)	0.00379*** (0.00142)	5.18e-10 (1.92e-08)
Constant	0.129 (0.0865)	23,486 (39,102)	1.238 (0.960)	0.0461** (0.0210)	-720.5 (5,735)	0.400* (0.234)
Observations	2,068	2,068	2,068	11,626	11,626	11,626
Wald (stdPG = augmPG) p-value	0.2540	0.2357	0.2915	0.5482	0.3388	0.4775

Notes: Table presents results for payment compliance outcomes by age cohort. Columns 1-3 present results for taxpayers 35 years or younger and columns 4-6 present results for taxpayers older than 35 years. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table 24 summarizes results for taxpayers with high and low PIT arrears.⁷⁹ Columns 1 – 3 present results for taxpayers with low PIT arrears and suggest that the standard and augmented public goods messages increase payment probability by 2.4 and 3.4 percentage points respectively. Results using the IHS transformation of the outcome variable indicate an increase in PIT paid by 25 and 35 percent respectively. However, results for the levels regression indicate an increase of J\$1, 328 (US\$10) from the augmented public goods message only. Results in columns 4 – 6 indicate positive compliance effects for the probability of payment and the IHS transformation, for the standard public good message only.

⁷⁹ We define low PIT arrears as being at or below J\$86, 046, and high as amounts in excess of this threshold. The cut off corresponds with the 75th percentile of the PIT arrears distribution.

Table 24: Effect of Public Goods Messages on Payment of PIT Arrears, by Level of PIT Arrears

	IIT Arrears ≤ 75th Percentile			IIT Arrears > 75th Percentile		
	Prob of Paying (1)	Amount Paid (2)	IHS Transformation (3)	Prob of Paying (4)	Amount Paid (5)	IHS Transformation (6)
Public Goods (stdPG)	0.0241*** (0.00923)	1,092 (787.4)	0.252** (0.0977)	0.0538*** (0.0173)	8,045 (6,800)	0.600*** (0.206)
Public Goods & Spending Info (augmPG)	0.0338*** (0.00954)	1,328* (775.5)	0.345*** (0.101)	0.0264 (0.0169)	6,321 (8,272)	0.281 (0.202)
Risk	-0.00986 (0.00844)	-385.6 (953.8)	-0.103 (0.0905)	0.0816*** (0.0109)	15,991*** (4,336)	1.005*** (0.130)
Age	-0.000410* (0.000245)	0.591 (22.95)	-0.00418 (0.00263)	-0.000357 (0.000456)	-139.0 (242.3)	-0.00390 (0.00542)
Single	-0.0105 (0.00662)	-879.3 (682.3)	-0.115 (0.0710)	-0.0344*** (0.0124)	-14,947*** (5,034)	-0.449*** (0.147)
Male	-0.00717 (0.00609)	-1,170* (635.6)	-0.0872 (0.0649)	0.0160 (0.0118)	-4,003 (5,400)	0.191 (0.143)
Email	0.174*** (0.0271)	11,241*** (3,131)	1.883*** (0.298)	0.195*** (0.0293)	34,408*** (11,163)	2.347*** (0.355)
PIT Arrears	-2.79e-07 (1.87e-07)	0.0645*** (0.0193)	-1.32e-06 (2.02e-06)	-7.28e-09*** (1.53e-09)	0.00176 (0.00131)	-8.03e-08*** (1.97e-08)
Constant	0.122*** (0.0248)	1,528 (2,943)	1.219*** (0.268)	-0.0305 (0.0332)	4,553 (15,531)	-0.429 (0.393)
Observations	10,272	10,272	10,272	3,422	3,422	3,422
Wald (stdPG = augmPG) p-value	0.4073	0.8180	0.4507	0.1952	0.8595	0.2047

Notes: Table presents results for payment compliance outcomes by level of PIT arrears. Columns 1-3 present results for taxpayers with PIT arrears less than or equal to the 75th percentile and columns 4-6 present results for those with PIT arrears greater than the 75th percentile. The 75th percentile of the PIT arrears distribution corresponds with arrears of J\$86,045. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Cost Benefit Analysis

We conduct a simple cost benefit analysis to assess the effectiveness of the messaging campaign. We focus on sub-experiment two, since the compliance effects in sub-experiment one were not statistically significant. Table C.2 highlight the main costs of the experiment, which include the cost of materials (paper, envelopes, toner, ink and glue), postage and labor. The total cost for sub experiment two is estimated at J\$1.2 million (US\$9,302). The cost of materials and stamps amounted to approximately J\$1.1 million (US\$8,527), roughly ninety percent of total cost. The experiment was ‘projectized’ by TAJ and eight staff members were ‘employed’ to print, stack and mail letters over seven days, from August 12 – 19. Total labor costs are estimated at approximately J\$119 thousand (US\$922) or roughly ten percent of total costs. These are direct costs and do not take into account, for example, work done by staff to provide data

support or costs associated with dealing with increased call traffic to the customer care help line. Table C.3 gives estimates of the costs and revenue gains from the standard and augmented public goods treatments. Average cost per letter is estimated at approximately J\$205 (US\$1.60). Estimated costs of sending letters to taxpayers selected to receive the standard public goods message is approximately J\$590,352 (US\$4,576) and the similar costs associated for the augmented public goods message is approximately J\$584,824 (US\$4,534).

Using our most conservative estimates, the standard public goods message increased payment of PIT arrears by an average of approximately J\$2,590 (US\$20). The total increase in revenue from this treatment is estimated at J\$7.4 million (US\$57,364); more than 12 times the cost of sending the standard public goods messages and about 6 times the cost of the entire sub-experiment. The additional revenue from the standard message is less than 0.1 percent of outstanding PIT arrears. If we assume the effects from the standard message apply across the entire sample, potential revenue gains are estimated at J\$35 million (US\$271,317), barely 0.5 percent of total PIT arrears. These results suggest that even though the additional revenues generated are not significant in nominal terms or as a share of total PIT arrears, when compared to the costs of sending the letters, the messages are still very cost effective.

Conclusion

This essay examined the effect of public goods messaging on PIT compliance with contemporaneous quarterly payment obligations and payment of PIT arrears. We extend the literature that uses messaging to elicit behavioral responses for taxpayers in developing countries by examining the incremental compliance effects of an augmented public goods message. The augmented message provided additional information to taxpayers which outlined actual spending

on key public services in the previous year. This was designed to improve their perception of transparency of government, to build trust and to induce a sense of taxpayer agency – their ability to influence how public resources are spent.

The messages had no effect on quarterly payments, for which there are no penalties for non-compliance, but increased payment of PIT arrears. Taxpayers receiving the standard public goods message are more likely to make a payment and pay larger amounts of outstanding PIT arrears relative to the control group. In general these results are in line with findings from essay two, which suggests the possibility of a complementarity in the relationship between the enforcement strength of the taxing regime and non-pecuniary factors. Relative to the standard public goods message, including information about public spending allocation did not have a differential effect on compliance. We find some evidence of heterogeneity, with older taxpayers responding to both the standard and augmented public goods messages, but no effect on younger taxpayers. We find only limited heterogeneous effects between high and low tax debtors – low tax debtors respond to both messages but high debtors respond only to the standard public goods message.

Our results point to the usefulness of public goods messaging to encourage payment of PIT arrears – an obligation that attracts penalties and interest, but also to the ineptness at encouraging compliance with quarterly payments – an obligation that is totally voluntary and attracts no penalty for non-compliance. One implication of this result is that to improve compliance with quarterly payments, in addition to non-pecuniary strategies, the tax administration should impose and enforce pecuniary penalties for non-compliance. This could

possibly be generalized to other instances where the tax administration is considering using non-pecuniary measures to boost compliance in a zero enforcement environment.

The results also provide guidance on who may be more likely to respond to these messages and also what variant of the message is likely to be most effective. Another important policy implication from our results is that messaging or communication strategies targeted at older cohorts may be more effective at boosting payments of tax arrears, compared to younger cohorts. Greater efforts are needed to engage younger cohorts in order to improve compliance. This may involve using new communication technologies – such as different delivery mechanisms and social media engagement. An alternative might be to structure compliance programs which target enforcement toward the young. These results are particularly useful for developing countries with limited resources to adequately police taxpayers. Utilizing simple and relatively cheap but effective communication strategies, targeted at taxpayers who are most likely to respond can provide a significant boost to revenues. Additionally, including these types of messaging in the compliance arsenal can potentially allow for more efficient deployment of scarce tax administration resources and can significantly boost operational efficiency within the tax administration.

CONCLUSION

This dissertation comprises three essays that examine critical aspects of fiscal policy and explores important determinants of tax compliance in a developing country context. We move from examining the fiscal response to changes in debt-to-GDP for developing countries in the first essay, to exploring possible determinants of tax compliance in Jamaica – a key factor that impacts a country’s ability to make the necessary fiscal adjustments, in essays two and three.

These essays make significant contributions to their respective literatures. In the first essay we highlight an omission in the literature, where the tendency has been to focus on the relationship between the primary balance and debt-to-GDP in more advanced economies. We attempt to plug this gap first by broadening this discussion to include developing countries. Further, we examine the relationship between a country’s debt-to-GDP and its ‘fiscal effort’ – a variable we argue is much more closely aligned to the fiscal policy discretion of governments. Next we dichotomize the fiscal reaction function (FRF) into its general revenue, tax revenue and primary spending components and examine each separately.

Essays two and three examine key factors that influence tax compliance in a developing country context. We are the first to employ quasi-experimental and experimental approaches to examine the effects of taxpayer services and public goods messaging on tax compliance in an English speaking LAC developing country. We examine compliance for a range of outcomes, for different tax types and across the large, medium and small taxpayer segments. Further we vary the enforcement strength of the taxing regime and examine the interaction with non-pecuniary factors, to determine the relationship between the two.

We find that developing countries increase their primary balance and ‘fiscal effort’ in response to debt and that these policy adjustments are done along both the revenue and

expenditure margins. We also find some suggestive evidence that countries make larger adjustments at higher levels of debt relative to lower levels of debt. We find more convincing evidence of a nonlinear response for the fiscal components. In particular, the countries examined, increase revenues when debt is ‘high’ but cut spending when debt is ‘low’. Tests for heterogeneity in fiscal response suggest that whereas higher and upper-middle income countries make significant adjustments along the revenue margin, low and lower-middle countries do not. This may be suggestive of revenue mobilization challenges facing the latter.

Low levels of tax compliance is at least one reason some developing countries find it difficult to grow revenues. Our results from the second and third essays suggest that non-pecuniary factors such as the provision of tax payer services and public goods messaging can potentially improve compliance in developing countries. In the context of Jamaica, taxpayer services improved compliance among large taxpayers, for the GCT but not the CIT. Public goods messaging were only effective in boosting payments of PIT arrears but not for quarterly PIT obligations. Notably, the positive compliance effects were for outcomes where the legal enforcement context was stronger relative to those outcomes where enforcement was weak. This provides suggestive evidence of a complementarity between the relative strength of the legal enforcement framework of the taxing regime and non-pecuniary determinants of compliance.

The results have important policy implications. One obvious implication of this is that developing countries must strengthen their revenue generating capacity in order to better sustain their debt. This is particularly true for countries with relatively high levels of debt. The extent to which countries are able to boost revenues will depend on the severity of the revenue mobilization challenges they face. A big part of the problem in many developing countries is from low levels of tax compliance.

Improving compliance will require the appropriate mix of enforcement and non-pecuniary or ‘softer’ interventions. Notably however, our findings caution the adoption of a softer approach to tax administration that attempt to encourage compliance by leveraging tax moral factors in a context of major enforcement deficiencies - example weak legal framework and corruption. A strong(er) legal enforcement framework is required if non-pecuniary factors such as taxpayer service provision and appeals to reciprocity are to be effective in improving compliance. Further, these interventions should be targeted at segments of the taxpayer population who are more likely to respond. Use of fiscal intermediaries or alternative compliance mechanisms such as industry oversight bodies and the imposition of tax compliance requirements to carry out primary economic activities, can serve as potential substitutes for provision of taxpayer services and personalized messaging strategies. Utilizing these ‘compliance intermediaries’ could provide a low cost solution to the already resource strapped tax administrations across developing countries.

Appendix A: Appendix for Essay One

Table A.1: List of Countries, Income Category

<p>High and Upper Middle Income Countries: Argentina, Belarus, Belize, Brazil Bulgaria, Chile, Colombia, Costa Rica, Dominica Dominican Republic, Fiji, Grenada, Israel, Jamaica, Jordan, Korea (South), Latvia, Lithuania, Malaysia, Mauritius, Mexico, Mongolia, Namibia, Paraguay, Peru, Poland, Romania, South Africa, Suriname, Thailand, Tunisia, Turkey, Venezuela</p> <p>Low and Lower Middle Income Countries: Bangladesh, Bolivia, Egypt, El Salvador, Ethiopia, Georgia, Ghana, Guatemala, Honduras, India, Indonesia, Kenya, Moldova, Morocco, Pakistan, Papua New Guinea, Philippines, Sri Lank, Uganda, Ukraine</p>
--

Notes: Country classifications done using the World Bank Analytical Classification based on the World Bank Atlas Method. Low income countries have per capita GNI less than or equal to \$1,025; Lower-middle income countries between \$1,206 - \$4,035, upper-middle income countries between \$4,036 - \$12,475 and high income countries above \$12,476.

Table A.2: Source and Description of Variables

Variable	Description	Source
Primary Balance	Consolidated central government revenue less primary expenditure	IMF, GFS
Fiscal Effort	Consolidated central government tax revenue less primary expenditure	IMF, GFS
Central Govt. Revenue	Consolidated central government revenues	IMF, GFS
Central Govt. Tax Revenue	Consolidated central government tax revenues	IMF, GFS
Central Govt. Primary Spending	Consolidated government primary spending	IMF, GFS
General Govt. Debt / GDP	One period lag of general government debt-to-GDP ratio	IMF Fiscal Affairs Department, Historical Public Debt Data
Output Gap	Output gap calculated as deviations from the Hedrick Prescott trend	World Bank, WDI
Expenditure Gap	Expenditure gap calculated as deviations from the Hedrick Prescott trend	IMF, WEO
Income Tax / Total Tax Revenue	Consolidated central government total income tax revenue as a share of total revenue	IMF, GFS
Inflation Rate	Based on the GDP deflator	World Bank, WDI
Openness	Sum of imports and exports as a share of GDP	World Bank, WDI
Net ODI Share in GNI	Official development assistance as a share of GNI	World Bank, WDI
Institutional Democracy	This index ranges from -10 to +10 ten, with the boundaries representing weakest and strongest democracy respectively.	Polity IV Data Base
IMF or Fiscal Rule	Dummy variable equal 1 if country has a fiscal rule or party to an IMF agreement	IMF - History of Lending Data and Fiscal Rule Data Bases
Agriculture Share in GDP	Agriculture value added as a share of GDP	World Bank, WDI

Notes: Table gives a description and data sources for the variables used in the analysis.

Table A. 3: Sensitivity Analysis Primary Balance, Full Sample

	Primary Balance			
	DGMM (1)	DGMM (2)	DGMM (3)	DGMM (4)
lag Dependent Variable	0.559*** (0.0723)	0.547*** (0.0684)	0.542*** (0.0657)	0.548*** (0.0681)
Lag Debt/GDP	0.0337*** (0.0102)	0.0329*** (0.00988)	0.0328*** (0.00967)	0.0318*** (0.00965)
Output Gap	0.0598** (0.0233)	0.0618** (0.0247)	0.0627** (0.0257)	0.0614** (0.0257)
Expenditure Gap	-0.0206 (0.0190)	-0.0187 (0.0194)	-0.0183 (0.0196)	-0.0160 (0.0197)
Inflation	0.000656** (0.000303)	0.000741*** (0.000278)	0.000799*** (0.000295)	0.000813*** (0.000299)
Trade Openness	0.00446 (0.00881)	0.00354 (0.00881)	0.00379 (0.00873)	0.00761 (0.0101)
Share of Agriculture in GDP	0.0241 (0.0455)	0.0260 (0.0446)	0.0311 (0.0440)	0.0338 (0.0424)
Share - Income Tax	-0.0142 (0.0154)	-0.0120 (0.0149)	-0.0125 (0.0150)	-0.0107 (0.0154)
Polity IV	-0.000951*** (0.000317)	-0.000927*** (0.000308)	-0.000940*** (0.000300)	-0.000925*** (0.000302)
Net ODI/GNI		0.116* (0.0617)	0.113* (0.0579)	0.111** (0.0549)
Fiscal Rule/IMF			0.00207 (0.00412)	0.00229 (0.00412)
IMF Program				0.00196 (0.00218)
Observations	588	588	588	567
Number of Countries	43	43	43	41
AR(2) (p-value)	0.261	0.245	0.243	0.136
Hansen J (p-value)	0.994	0.993	0.995	0.998
Year Dummies	yes	yes	yes	yes

Notes: Dependent variable is the primary balance expressed as a share of GDP. Results are for difference GMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995) with the lag dependent variable and the lag debt/GDP assumed to be potentially endogenous. To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables as instruments. AR (2) diagnostic tests suggest no second order serial correlation in the errors. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A.4 Sensitivity Analysis Fiscal Effort, Full Sample

	Fiscal Effort			
	DGMM (1)	DGMM (2)	DGMM (3)	DGMM (4)
lag Dependent Variable	0.545*** (0.148)	0.547*** (0.148)	0.548*** (0.147)	0.547*** (0.151)
Lag Debt/GDP	0.0226** (0.0109)	0.0234** (0.0112)	0.0229** (0.0108)	0.0223** (0.0109)
Output Gap	0.0610** (0.0296)	0.0602** (0.0289)	0.0612** (0.0277)	0.0585** (0.0268)
Expenditure Gap	-0.0258 (0.0160)	-0.0266* (0.0158)	-0.0263* (0.0158)	-0.0228 (0.0155)
Inflation	-0.00104*** (0.000390)	-0.00108*** (0.000406)	-0.000979** (0.000420)	-0.000974** (0.000443)
Trade Openness	0.00493 (0.0113)	0.00551 (0.0111)	0.00598 (0.0106)	0.00776 (0.0118)
Share of Agriculture in GDP	0.0774 (0.0564)	0.0757 (0.0561)	0.0851 (0.0566)	0.0912 (0.0558)
Share - Income Tax	0.0490 (0.0309)	0.0475 (0.0312)	0.0464 (0.0311)	0.0519 (0.0338)
Polity IV	-0.000574* (0.000299)	-0.000594* (0.000306)	-0.000607** (0.000296)	-0.000576* (0.000297)
Net ODI/GNI		-0.0707 (0.0609)	-0.0750 (0.0584)	-0.0765 (0.0598)
Fiscal Rule/IMF			0.00393 (0.00427)	0.00407 (0.00435)
IMF Program				0.00236 (0.00241)
Observations	588	588	588	567
Number of Countries	43	43	43	41
AR(2) (p-value)	0.234	0.230	0.241	0.173
Hansen J (p-value)	0.999	0.999	0.999	0.999
Year Dummies	yes	yes	yes	yes

Notes: Dependent variable is the fiscal effort expressed as a share of GDP. Results are for difference GMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995) with the lag dependent variable and the lag debt/GDP assumed to be potentially endogenous. To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables as instruments. AR (2) diagnostic tests suggest no second order serial correlation in the errors. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A.5: Sensitivity Analysis General Revenue, Full Sample

	General Revenue			
	DGMM (1)	DGMM (2)	DGMM (3)	DGMM (4)
lag Dependent Variable	0.479*** (0.163)	0.468*** (0.159)	0.457*** (0.165)	0.467*** (0.162)
Lag Debt/GDP	0.0272** (0.0131)	0.0247** (0.0125)	0.0257** (0.0129)	0.0229* (0.0124)
Output Gap	0.0296 (0.0343)	0.0277 (0.0342)	0.0268 (0.0339)	0.0311 (0.0342)
Expenditure Gap	0.0143 (0.0191)	0.0180 (0.0193)	0.0190 (0.0195)	0.0153 (0.0197)
Inflation	0.000568 (0.000642)	0.000669 (0.000626)	0.000604 (0.000617)	0.000640 (0.000599)
Trade Openness	0.0190 (0.0121)	0.0175 (0.0119)	0.0174 (0.0116)	0.0223* (0.0129)
Share of Agriculture in GDP	-0.0654 (0.0632)	-0.0553 (0.0608)	-0.0610 (0.0614)	-0.0577 (0.0600)
Share - Income Tax	-0.0415 (0.0337)	-0.0380 (0.0337)	-0.0374 (0.0337)	-0.0397 (0.0344)
Polity IV	-0.000589 (0.000473)	-0.000559 (0.000470)	-0.000565 (0.000480)	-0.000596 (0.000472)
Net ODI/GNI		0.177* (0.102)	0.178* (0.105)	0.163 (0.106)
Fiscal Rule/IMF			-0.00292 (0.00549)	-0.00223 (0.00561)
IMF Program				0.00239 (0.00247)
Observations	612	612	612	585
Number of Countries	43	43	43	41
AR(2) (p-value)	0.518	0.479	0.402	0.444
Hansen J (p-value)	0.980	0.999	0.999	1
Year Dummies	yes	yes	yes	yes

Notes: Dependent variable is general revenues expressed as a share of GDP. Results are for difference GMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995) with the lag dependent variable and the lag debt/GDP assumed to be potentially endogenous. To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables as instruments. AR (2) diagnostic tests suggest no second order serial correlation in the errors. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A.6: Sensitivity Analysis Tax Revenues, Full Sample

	Tax Revenue			
	DGMM (1)	DGMM (2)	DGMM (3)	DGMM (4)
lag Dependent Variable	0.554*** (0.0967)	0.557*** (0.0964)	0.556*** (0.0970)	0.556*** (0.0878)
Lag Debt/GDP	0.0143** (0.00602)	0.0144** (0.00576)	0.0145** (0.00594)	0.0115** (0.00531)
Output Gap	0.0232 (0.0216)	0.0230 (0.0213)	0.0229 (0.0213)	0.0229 (0.0217)
Expenditure Gap	0.0139 (0.0181)	0.0135 (0.0183)	0.0135 (0.0182)	0.0108 (0.0186)
Inflation	-0.00117** (0.000485)	-0.00117** (0.000486)	-0.00118** (0.000494)	-0.00117** (0.000485)
Trade Openness	0.0155** (0.00761)	0.0156** (0.00763)	0.0156** (0.00757)	0.0187** (0.00868)
Share of Agriculture in GDP	0.00293 (0.0337)	0.00198 (0.0352)	0.00151 (0.0367)	0.00740 (0.0349)
Share - Income Tax	0.0217 (0.0164)	0.0212 (0.0162)	0.0213 (0.0162)	0.0224 (0.0166)
Polity IV	-0.000136 (0.000324)	-0.000140 (0.000323)	-0.000140 (0.000324)	-0.000150 (0.000321)
Net ODI/GNI		-0.0202 (0.0628)	-0.0201 (0.0629)	-0.0260 (0.0613)
Fiscal Rule/IMF			-0.000340 (0.00307)	0.000111 (0.00309)
IMF Program				0.00200 (0.00169)
Observations	611	611	611	584
Number of Countries	43	43	43	41
AR(2) (p-value)	0.664	0.673	0.674	0.749
Hansen J (p-value)	0.998	0.998	0.998	1
Year Dummies	yes	yes	yes	yes

Notes: Dependent variable is tax revenue expressed as a share of GDP. Results are for difference GMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995) with the lag dependent variable and the lag debt/GDP assumed to be potentially endogenous. To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables as instruments. AR (2) diagnostic tests suggest no second order serial correlation in the errors. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A.7: Sensitivity Analysis Primary Spending, Full Sample

	Primary Spending			
	DGMM (1)	DGMM (2)	DGMM (3)	DGMM (4)
lag Dependent Variable	0.596*** (0.105)	0.594*** (0.106)	0.588*** (0.112)	0.546*** (0.117)
Lag Debt/GDP	-0.0163* (0.00857)	-0.0178* (0.00917)	-0.0172** (0.00875)	-0.0168* (0.00978)
Output Gap	-0.0339 (0.0304)	-0.0339 (0.0302)	-0.0358 (0.0292)	-0.0356 (0.0306)
Expenditure Gap	0.0286* (0.0166)	0.0291* (0.0161)	0.0289* (0.0160)	0.0272 (0.0170)
Inflation	-6.54e-05 (0.000425)	-1.80e-05 (0.000425)	-0.000130 (0.000444)	-0.000152 (0.000430)
Trade Openness	0.0102 (0.0116)	0.00967 (0.0114)	0.00927 (0.0108)	0.0116 (0.0123)
Share of Agriculture in GDP	-0.0770* (0.0459)	-0.0749* (0.0448)	-0.0858** (0.0417)	-0.0918** (0.0445)
Share - Income Tax	-0.0265 (0.0282)	-0.0249 (0.0286)	-0.0237 (0.0280)	-0.0274 (0.0303)
Polity IV	0.000444 (0.000387)	0.000471 (0.000394)	0.000489 (0.000397)	0.000461 (0.000426)
Net ODI/GNI		0.0843 (0.0701)	0.0895 (0.0700)	0.0834 (0.0751)
Fiscal Rule/IMF			-0.00435 (0.00454)	-0.00468 (0.00489)
IMF Program				-0.000142 (0.00195)
Observations	588	588	588	567
Number of Countries	43	43	43	41
AR(2) (p-value)	0.927	0.932	0.979	0.982
Hansen J (p-value)	1	1	1	1
Year Dummies	yes	yes	yes	yes

Notes: Dependent variable is primary spending expressed as a share of GDP. Results are for difference GMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995) with the lag dependent variable and the lag debt/GDP assumed to be potentially endogenous. To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables as instruments. AR (2) diagnostic tests suggest no second order serial correlation in the errors. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table A.8: Response of the Primary Balance and Fiscal Effort for Sub-periods

	Primary Balance			Fiscal Effort		
	DGMM	DGMM	DGMM	DGMM	DGMM	DGMM
	1990-2011	1990-2008	1990-2000	1990-2011	1990-2008	1990-2000
	(1)	(2)	(3)	(4)	(5)	(6)
Lag Dependent Variable	0.559*** (0.0723)	0.598*** (0.0712)	0.618*** (0.128)	0.545*** (0.148)	0.555*** (0.191)	0.670*** (0.168)
Lag Debt/GDP	0.0337*** (0.0102)	0.0362*** (0.0106)	0.0519*** (0.0131)	0.0226** (0.0109)	0.0258** (0.0122)	0.0426*** (0.0118)
Output Gap	0.0598** (0.0233)	0.0465* (0.0240)	0.0751* (0.0422)	0.0610** (0.0296)	0.0458 (0.0326)	0.112** (0.0465)
Expenditure Gap	-0.0206 (0.0190)	-0.00672 (0.0191)	-0.0272 (0.0212)	-0.0258 (0.0160)	-0.0136 (0.0178)	-0.0264 (0.0220)
Inflation	0.000656** (0.000303)	0.000694** (0.000322)	0.00104*** (0.000368)	-0.00104*** (0.000390)	-0.00104** (0.000478)	-0.000581 (0.000440)
Trade Openness	0.00446 (0.00881)	0.00530 (0.0109)	-0.00257 (0.0167)	0.00493 (0.0113)	0.00969 (0.0128)	-0.00581 (0.0208)
Share of Agriculture in GDP	0.0241 (0.0455)	0.0209 (0.0432)	0.0735 (0.0784)	0.0774 (0.0564)	0.0721 (0.0559)	0.134 (0.0935)
Share - Income Tax	-0.0142 (0.0154)	-0.0267 (0.0195)	-0.0389** (0.0165)	0.0490 (0.0309)	0.0388 (0.0338)	-0.0122 (0.0172)
Polity IV	-0.000951*** (0.000317)	-0.000857*** (0.000326)	-0.00133*** (0.000471)	-0.000574* (0.000299)	-0.000530* (0.000308)	-0.00100** (0.000457)
Observations	588	489	225	588	489	225
No. Countries	43	43	33	43	43	33
AR(2) (p-value)	0.261	0.668	0.127	0.234	0.700	0.390
Hansen J (p-value)	0.994	0.938	0.307	0.999	0.985	0.530
Year Dummies	yes	yes	yes	yes	yes	yes

Notes: Table presents results for the primary balance and fiscal effort, expressed as a share of GDP, for the full sample period (1990-2011) and two sub periods 1990-2008 and 1990-2000. Results are for difference GMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995) with the lag dependent variable and the lag debt/GDP assumed to be potentially endogenous. To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the potentially endogenous variables as instruments. AR (2) diagnostic tests suggest no second order serial correlation in the errors. Robust standard errors in parentheses.*p<0.10, **p<0.05, ***p<0.01.

Table A.9: Response of Total Revenue, Tax Revenue and Primary Spending for Sub-periods

	Total Revenue			Tax Revenue			Primary Spending		
	DGMM	DGMM	DGMM	DGMM	DGMM	DGMM	DGMM	DGMM	DGMM
	1990-2011	1990-2008	1990-2000	1990-2011	1990-2008	1990-2000	1990-2011	1990-2008	1990-2000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lag Dependent Variable	0.479*** (0.163)	0.447** (0.210)	0.630*** (0.208)	0.554*** (0.0967)	0.607*** (0.116)	0.608*** (0.193)	0.685*** (0.0753)	0.728*** (0.0810)	0.751*** (0.0864)
Lag Debt/GDP	0.0272** (0.0131)	0.0286* (0.0163)	0.0439** (0.0222)	0.0143** (0.00602)	0.0146** (0.00598)	0.0313*** (0.00806)	-0.0127* (0.00698)	-0.0201** (0.00813)	-0.0215** (0.00892)
Output Gap	0.0296 (0.0343)	0.0415 (0.0399)	0.0367 (0.0358)	0.0232 (0.0216)	0.0332 (0.0226)	0.0671** (0.0338)	-0.0215 (0.0274)	0.00163 (0.0285)	-0.0254 (0.0391)
Expenditure Gap	0.0143 (0.0191)	0.0203 (0.0199)	0.0194 (0.0221)	0.0139 (0.0181)	0.0137 (0.0197)	0.0271 (0.0261)	0.0291* (0.0169)	0.0112 (0.0178)	0.0282 (0.0250)
Inflation	0.000568 (0.000642)	0.000705 (0.000694)	0.000717 (0.000738)	-0.00117** (0.000485)	-0.00110** (0.000543)	-0.000902** (0.000401)	-1.79e-06 (0.000457)	1.31e-05 (0.000503)	-0.000136 (0.000638)
Trade Openness	0.0190 (0.0121)	0.0161 (0.0148)	0.0171 (0.0137)	0.0155** (0.00761)	0.0144 (0.00911)	0.00792 (0.0130)	0.00830 (0.0102)	0.00538 (0.0108)	0.0205 (0.0157)
Share of Agriculture in GDP	-0.0654 (0.0632)	-0.0447 (0.0691)	-0.00407 (0.0613)	0.00293 (0.0337)	0.0192 (0.0335)	0.0842 (0.0613)	-0.0784* (0.0407)	-0.0602 (0.0400)	-0.0909 (0.0958)
Share - Income Tax	-0.0415 (0.0337)	-0.0467 (0.0353)	-0.0412* (0.0213)	0.0217 (0.0164)	0.0116 (0.0163)	-0.00669 (0.0179)	-0.0269 (0.0272)	-0.0306 (0.0297)	-0.00537 (0.0205)
Polity IV	-0.000589 (0.000473)	-0.000617 (0.000496)	-0.000260 (0.000355)	-0.000136 (0.000324)	-0.000182 (0.000317)	4.25e-05 (0.000328)	0.000407 (0.000333)	0.000278 (0.000351)	0.00104*** (0.000375)
Observations	612	513	242	611	512	241	588	489	225
No. Countries	43	43	36	43	43	36	43	43	33
AR(2) (p-value)	0.518	0.386	0.332	0.664	0.505	0.467	0.892	0.257	0.202
Hansen J (p-value)	0.980	0.761	0.336	0.998	0.973	0.687	1	0.988	0.364
Year Dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: Table presents results for general revenues, tax revenue and primary spending, all as a share of GDP, for the full sample period (1990-2011) and two sub periods 1990-2008 and 1990-2000. Results are for difference GMM regressions using forward orthogonal deviations (FOD) (Arellano and Bover, 1995) with the lag dependent variable and the lag debt/GDP assumed to be potentially endogenous. To minimize the problem of instrument proliferation, GMM instruments are constructed using only the second lag of the endogenous variables as instruments. AR (2) diagnostic tests suggest no second order serial correlation in the errors. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Figure A.1: Scatter of Fiscal Outcomes against Lagged debt-to-GDP, Full Sample

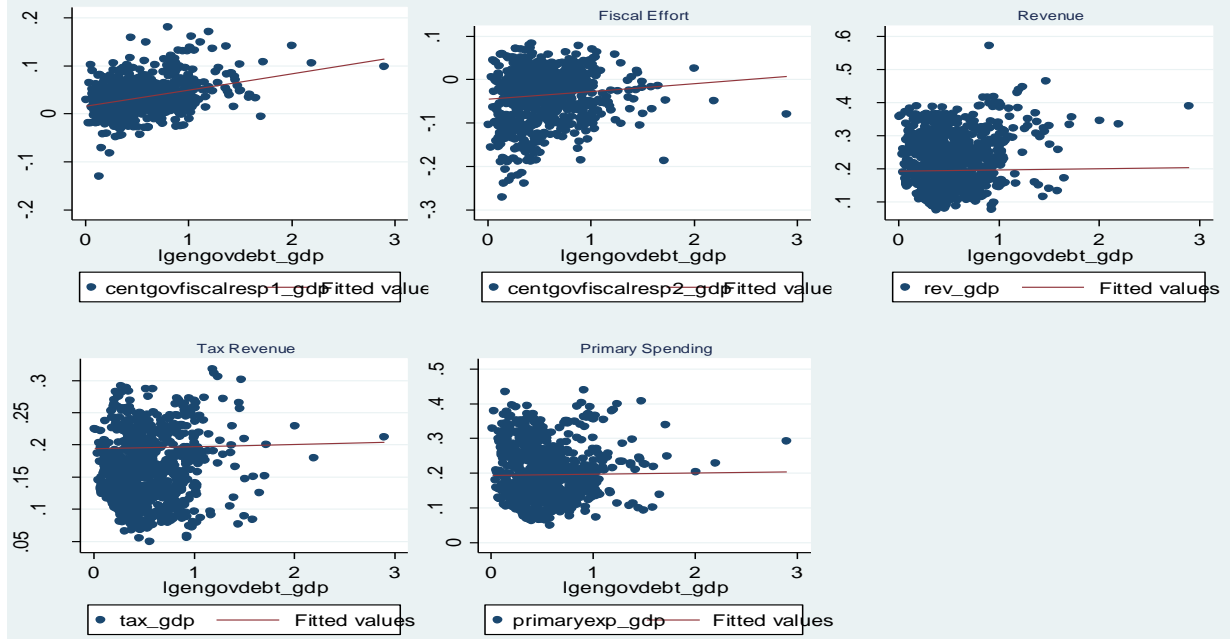


Figure A.2: Scatter of Fiscal Outcomes against Lagged debt-to-GDP for Low / Lower-middle Income Countries

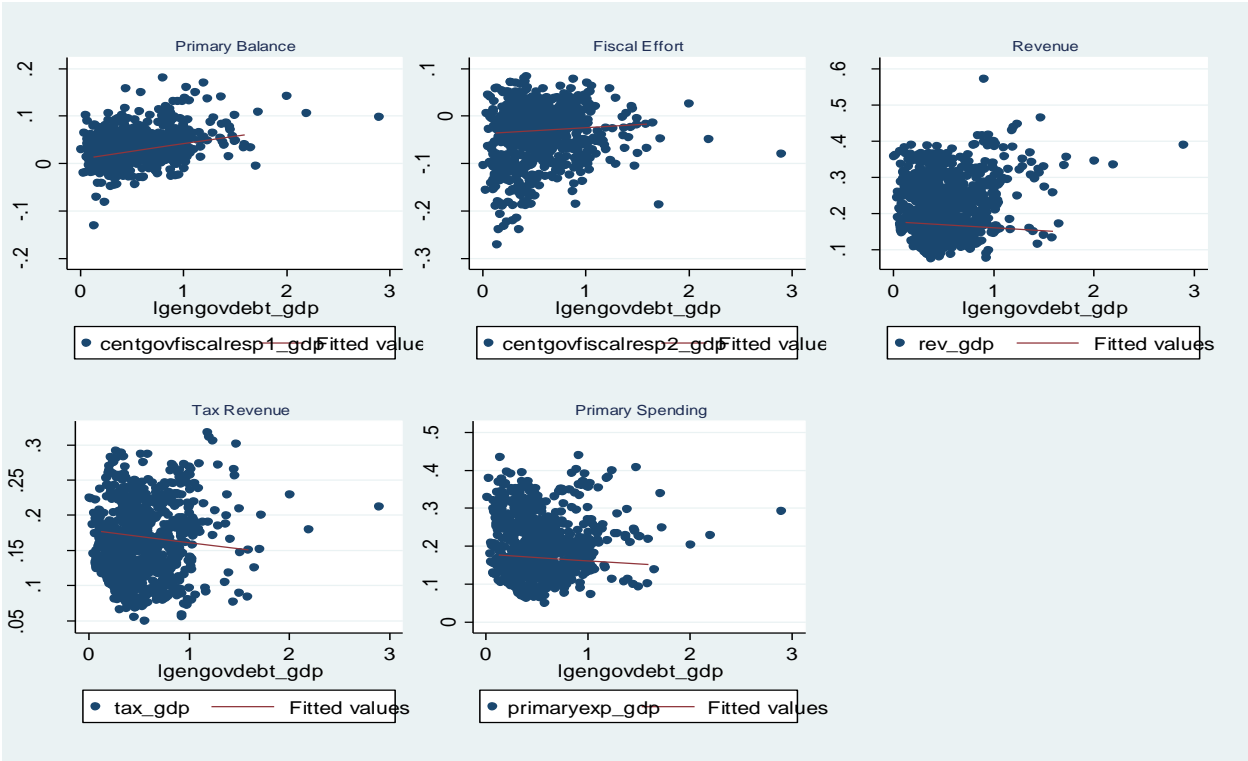
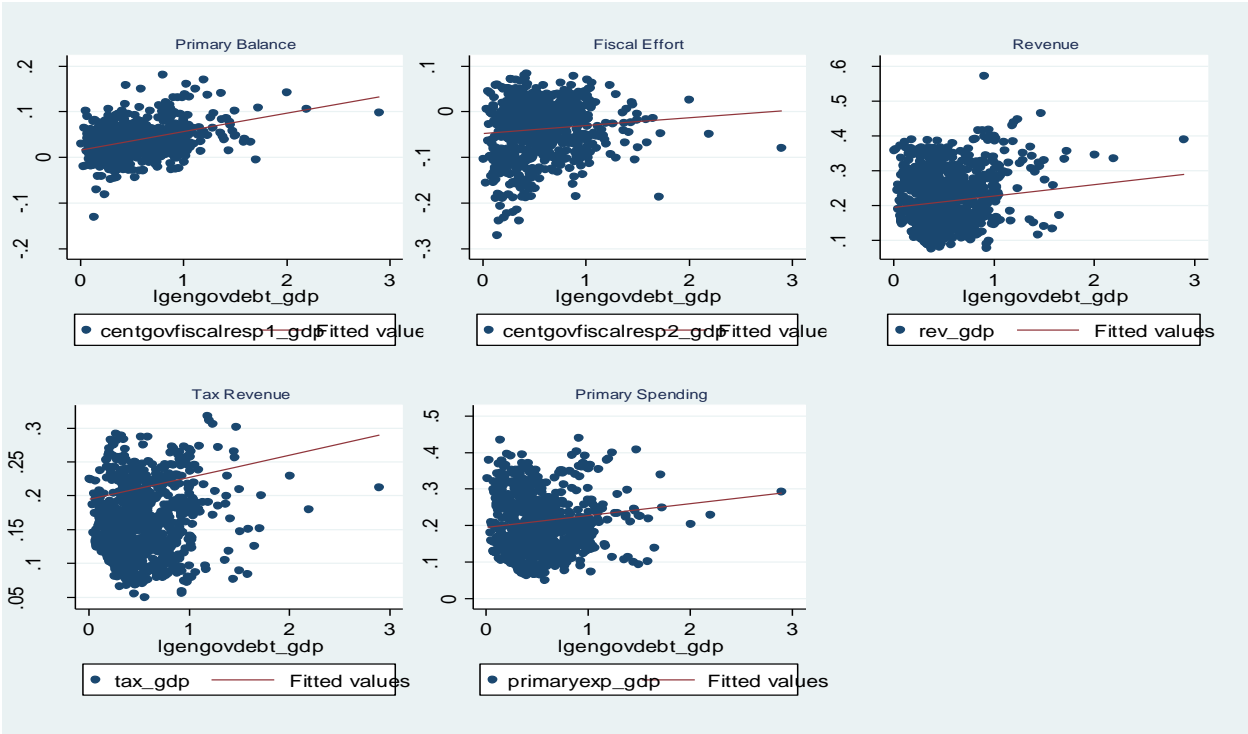


Figure A.3: Scatter of Fiscal Outcomes against Lagged debt-to-GDP, High / Upper-middle Income Countries



Appendix B: Appendix for Essay Two

Table B.1: Filing Compliance for All Taxpayers, Gross Receipts J\$1 million – J\$1 billion

	CIT		GCT	
	Prob Late	Days Late	Prob Late	Days Late
	(1)	(2)	(3)	(4)
LTO	0.472 (1.719)	5.904 (71.33)	0.509 (0.486)	8.343 (17.91)
Bandwidth (J\$ millions)	229	398	101	246
Observations	1,564	4,546	7,666	21,844
Year Dummy	no	no	no	no
Controls	no	no	no	no

Notes: This table presents estimates for filing compliance for taxpayers with reported gross receipts between J\$100 million and J\$1 billion and includes taxpayers who report both positive and zero tax liability. Columns 1–2 present results for the CIT and columns 3–4 present results for the GCT. Results are for the optimal bandwidth (IK). Regressions adopt a linear specification. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table B.2: Filing and Payment Compliance for Taxpayers who filed or paid both CIT and GCT

	No of Days		No of Days		Amount	Share
	Filed Late	Filed Late	Paid Late	Paid Late	Paid on Time	Paid on Time
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A - CIT Compliance</i>						
LTO	-2.882 (9.585)	-489.0 (773.0)	-2.010 (2.639)	-1,585 (1,990)	-3.287 (12.92)	1.543 (2.093)
IK Optimal Bandwidth (J\$ Million)	240 749 no no	283 930 no no	295 726 no no	306 766 no no	480 1,395 no no	303 756 no no
<i>Panel B - GCT Compliance</i>						
Panel B						
LTO	-0.255*** (0.0588)	-3.124*** (1.120)	-0.231*** (0.0636)	307.3*** (73.23)	3.845*** (0.995)	0.257*** (0.0726)
IK Optimal Bandwidth (J\$ Million)	130 4,389 no no	320 13,602 no no	152 5,236 no no	196 6,935 no no	122 4,046 no no	196 6,918 no no

Notes: This table presents estimates of filing and payment compliance for taxpayers who filed and paid both CIT and GCT for tax years 2009 – 2012. The columns show the outcome variables: probability of filing late, number of days filed late, probability of paying late, number of days paid late, amount paid on time and share of reported taxes paid on time. All estimates are based on their respective optimal bandwidths (IK). Regressions adopt a linear specification. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table B.3: GCT Filing and Payment Compliance for Financial and Non-Financial Sector

	Filed Late (1)	Days File Late (2)	Paid Late (3)	Days Paid Late (4)	Amount Paid OT (5)	Share Paid OT (6)
Financial Sector Taxpayers						
LTO	0.0699 (0.0670)	42.06 (122.8)	0.0761 (0.105)	-136.8 (200.3)	-53.57 (61.00)	-1.100 (1.047)
Observations	712	1,176	814	1,888	1,014	835
Year Dummy	no	no	no	no	no	no
Controls	no	no	no	no	no	no
Non Financial Taxpayers						
LTO	-0.172** (0.0694)	-15.34*** (5.748)	-0.145* (0.0758)	-258.1*** (51.48)	2.398*** (0.545)	0.157* (0.0908)
Observations	5,314	11,614	6,877	23,805	9336	7,246
Year Dummy	no	no	no	no	no	no
Controls	no	no	no	no	no	no
Bandwidth (J\$ Millions)	125	235	156	361	207	166

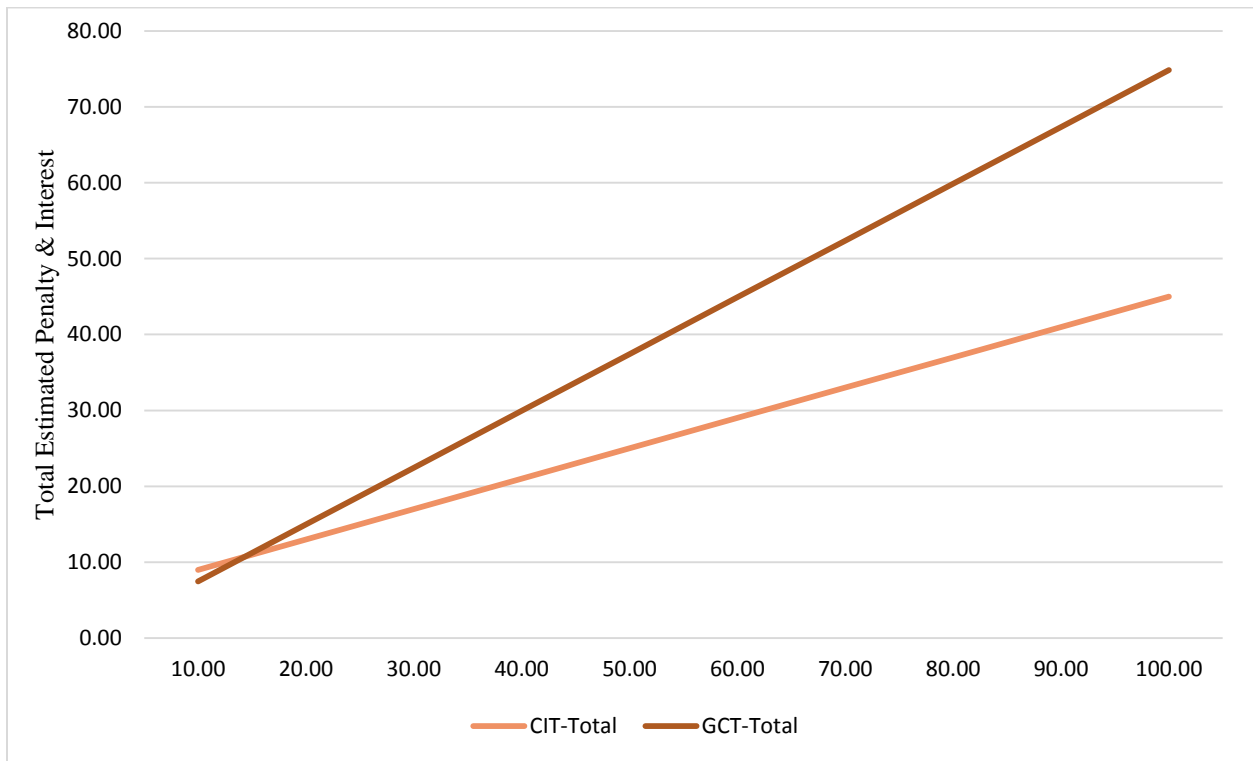
Notes: This table presents estimates of filing and payment compliance for ‘financial’ and ‘non-financial’ sector taxpayers. Taxpayers are linked to the financial and non-financial sectors using industry codes from the tax returns. All results presented are for the optimal bandwidth proposed by Imbens and Kalyanaraman (2009). Regressions adopt a linear specification. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Table B.4: GCT Filing and Payment Compliance for Importers and Non-Importers

	Filed Late (1)	Days File Late (2)	Paid Late (3)	Days Paid Late (4)	Amount Paid OT (5)	Share Paid OT (6)
Importers						
LTO	-0.227 (0.790)	-0.732 (1.535)	-0.102 (0.600)	106.4 (99.82)	3.897 (4.629)	-1.730 (4.162)
Observations	1,049	2,250	1,341	4,143	1,799	1,408
Year Dummy	no	no	no	no	no	no
Controls	no	no	no	no	no	no
Non-Importers						
LTO	-0.228** (0.0900)	-14.50** (5.775)	-0.208*** (0.0718)	-274.7*** (52.14)	4.343*** (0.806)	0.312** (0.134)
Observations	4,977	10,540	6,347	21,550	8,551	6,673
Year Dummy	no	no	no	no	no	no
Controls	no	no	no	no	no	no
Bandwidth (J\$ Millions)	125	235	156	361	207	166

Notes: This table presents estimates of filing and payment compliance for ‘importers’ and ‘non-importers’. We classify taxpayers as importers if their import / output ratio is greater than or equal to 20 percent receipts and taxpayer with a ratio less than 20 percent as non-importers. All results presented are for the optimal bandwidth proposed by Imbens and Kalyanaraman (2009). Regressions adopt a linear specification. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

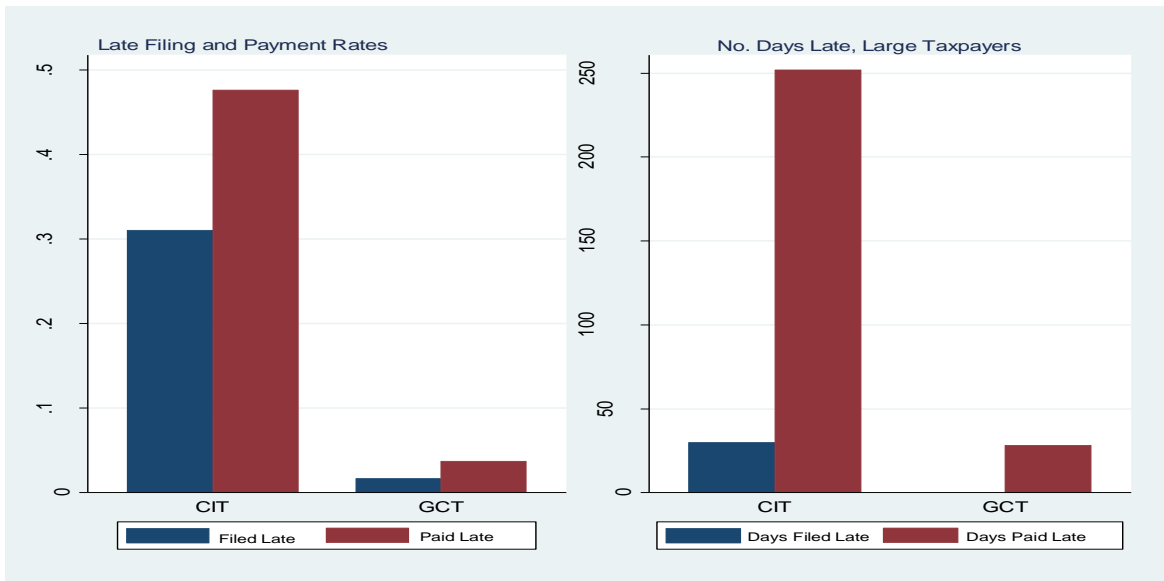
Figure B.1: Annual Estimated Penalty and Interest for Unpaid Tax Liability for CIT and GCT



Notes: Figure shows estimated penalty and interest (y axis) for unpaid and unfiled taxes (x axis) for the CIT and GCT, for a 12 month period. Estimates are based on the statutory penalty and interest rates for the respective taxes. Estimates do not include surcharges or any additional charge that taxpayers may incur and thus represents a lower bound estimate

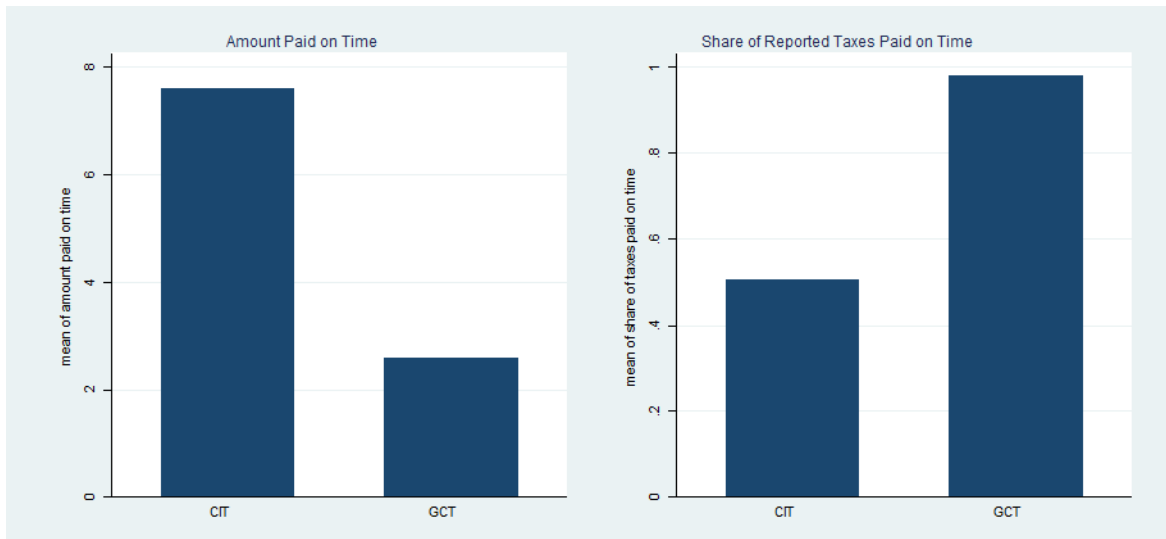
Figure B.2: Filing and Payment Compliance Patterns for Large Taxpayers

Panel A: CIT and GCT Filing and Payment Compliance Rates



Notes: Figure shows filing and payment compliance for taxpayers with reported gross receipts between J\$500 million and J\$1 billion who filed or paid both CIT and GCT for tax years 2009 – 2012. Column (1) shows the rate of late filing and payment for the CIT and GCT. Column (2) shows the number of days late that CIT and GCT is filed and paid.

Panel B: CIT and GCT Payments



Notes: Figure shows payment compliance for taxpayers with reported gross receipts between J\$500 million and J\$1 billion who filed or paid both CIT and GCT for tax years 2009 – 2012. Column (1) shows the amount of CIT and GCT paid on time and column (2) shows the share of reported CIT and GCT paid on time.

Figure B.3: Density Plots for CIT and GCT

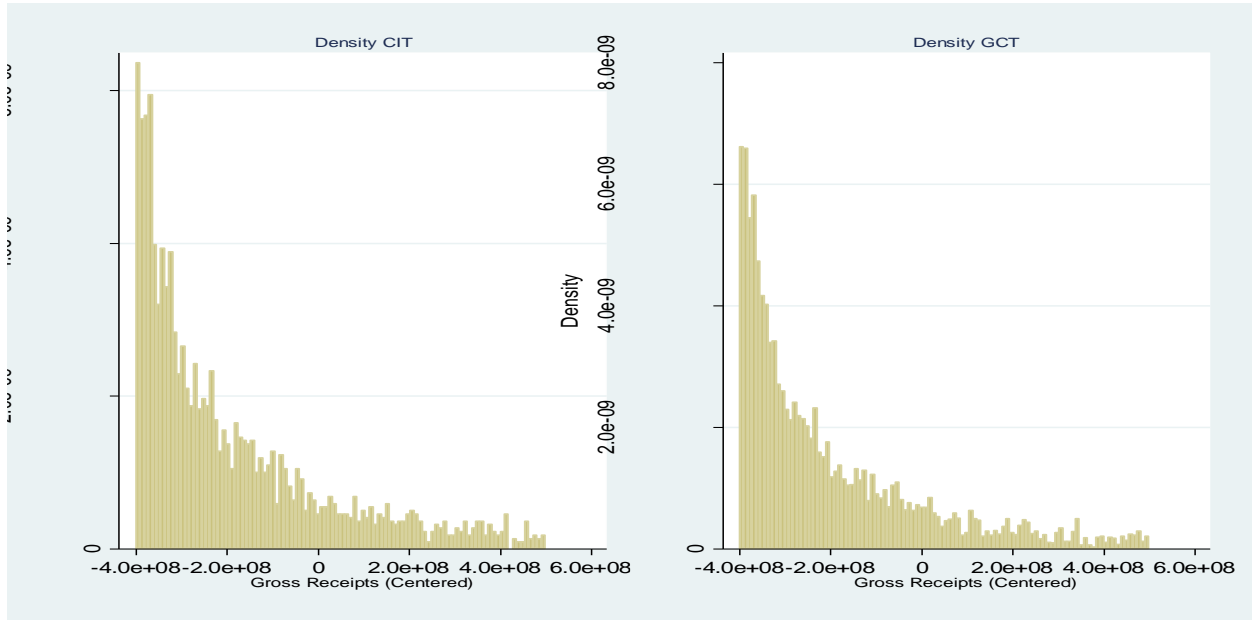


Figure B.4: McCrary's Test of Density Manipulation around the Threshold for the CIT and GCT

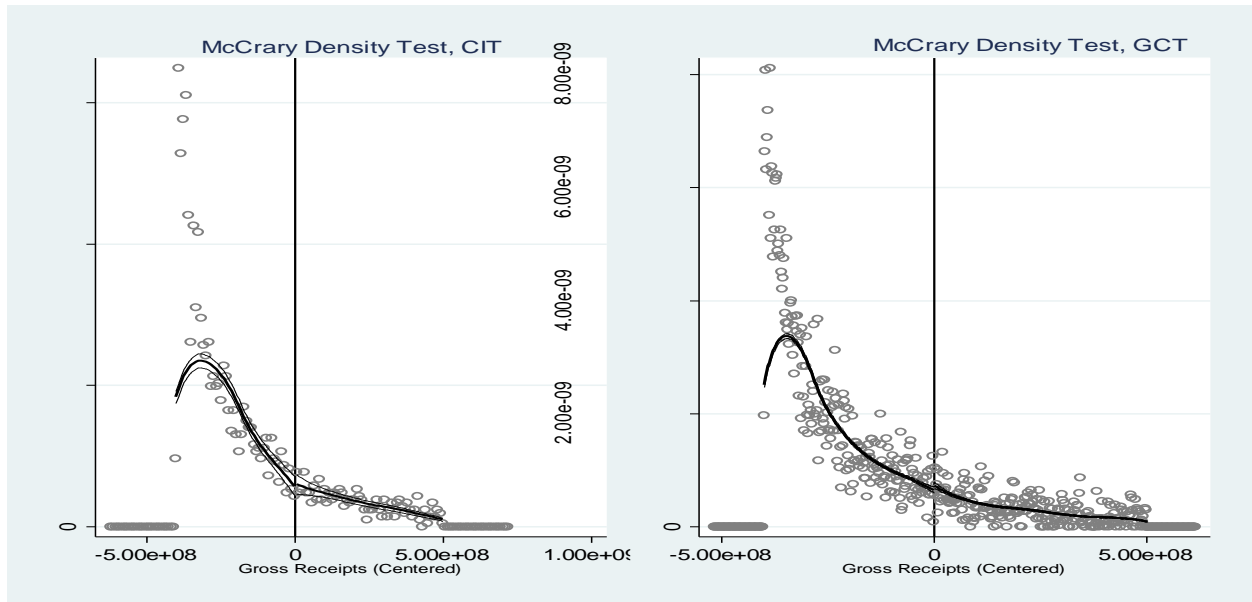
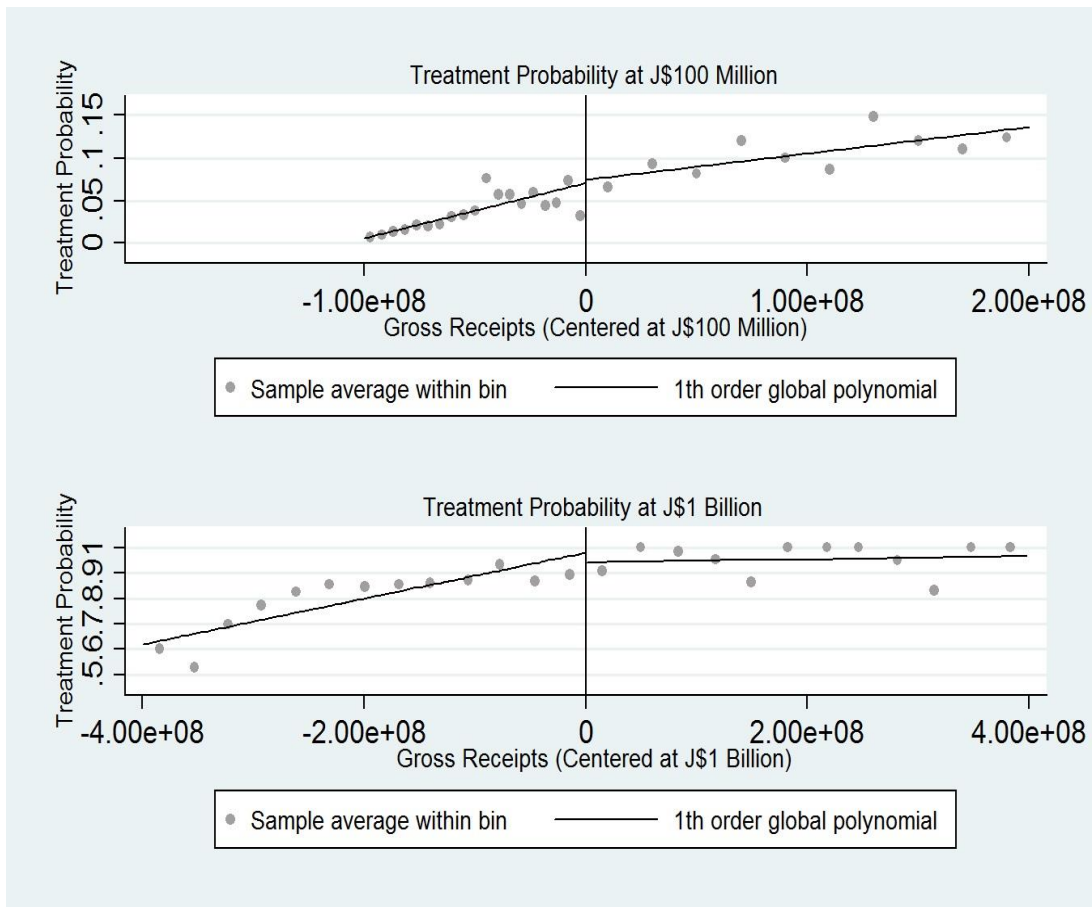
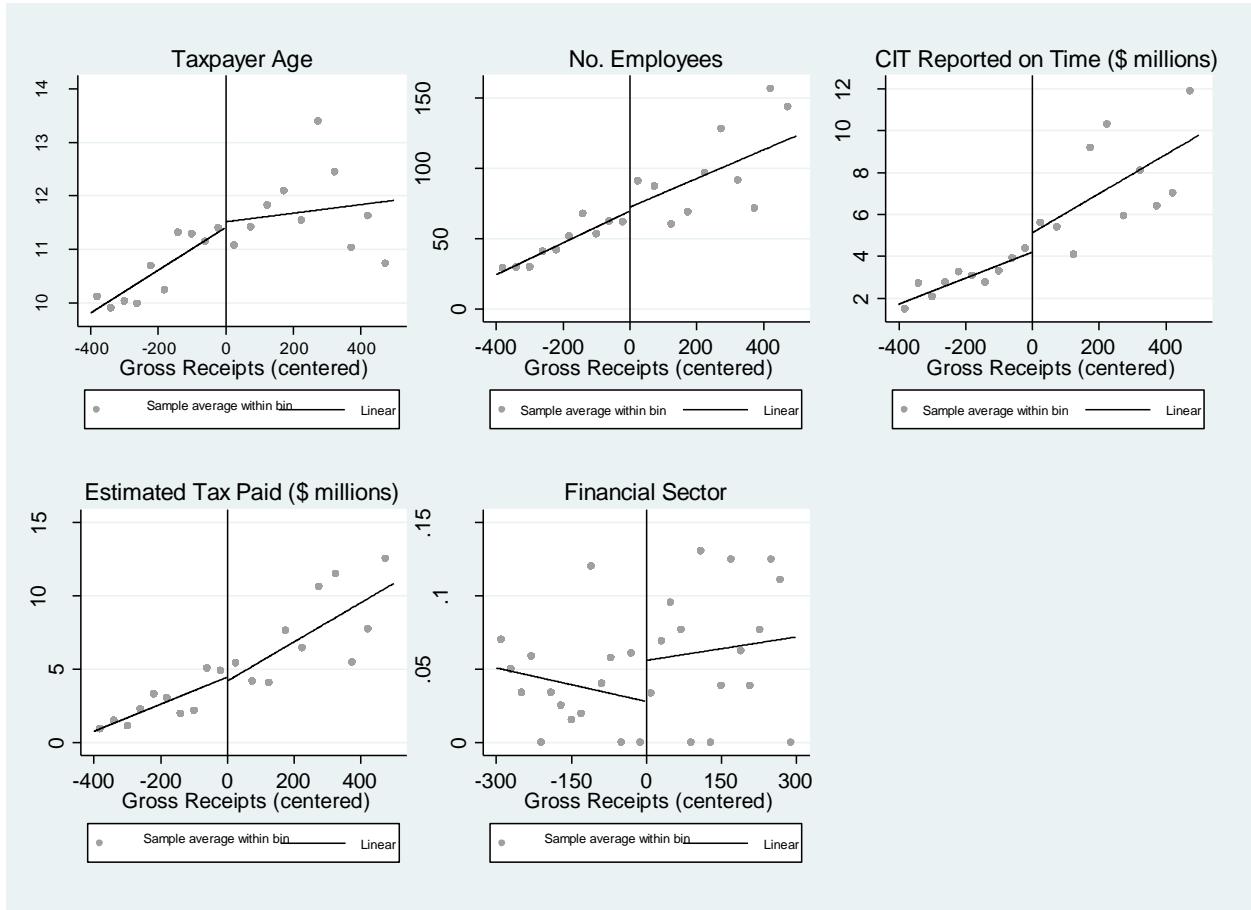


Figure B.5: Test for Discontinuity in Treatment Probability Placebo Thresholds



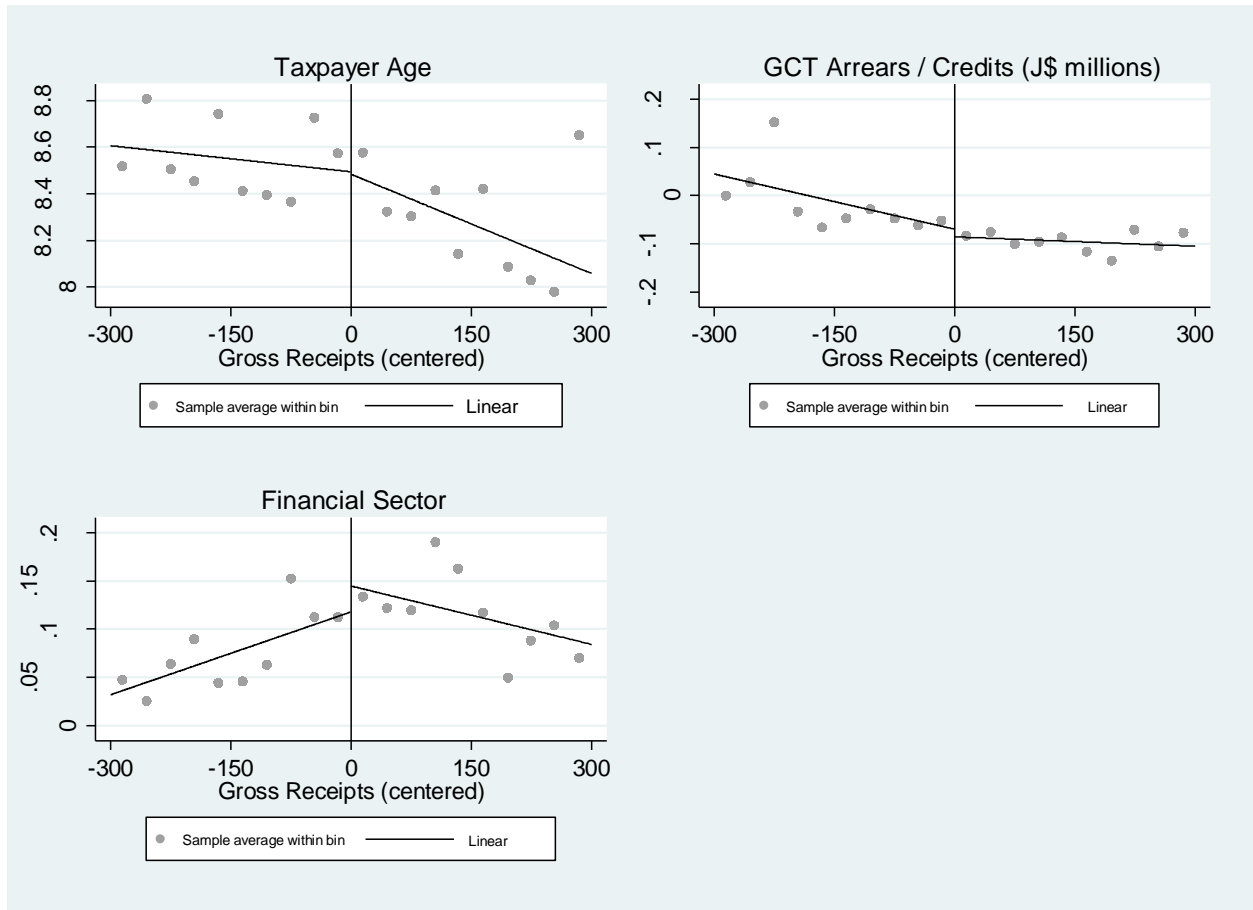
Notes: The placebo thresholds of J\$100 million and J\$1 billion were chosen to match TAJ's classification schedule for 'small' and 'large' taxpayers. TAJ classifies taxpayers with gross receipts less than or equal to J\$100 million as small and those with gross receipts of J\$1 billion or more as 'large'.

Figure B.6: Treatment Probability for Key Economic Sector for the CIT



Notes: RD plots done using integrated mean squared-error (IMSE) – equally spaced (ES) method with spacing estimators (Calonico et al., 2014a). The approach fits a local linear regression that approximates the conditional mean of the outcome variable (firm characteristics) to the left and right of the cut off.

Figure B.7: Tests for Discontinuity in Taxpayer Characteristics for the GCT



Notes: RD plots done using integrated mean squared-error (IMSE) – evenly spaced (ES) method with spacing estimators (Calonico et al.; 2014a). The approach fits a local linear regression that approximates the conditional mean of the outcome variable (firm characteristics) to the left and right of the cut off.

Appendix C: Appendix for Essay Three

Table C.1: Experimental Timeline

July 18 - 23, 2016	Data collection & sample selection
Collected data on all registered self employed taxpayers from TAJ Randomized the sample in treatment control groups for the two experiment in phase one	
August 3 - 5, 2016	Send letters for sub-experiment one
Letters for sub-experiment one were generated, printed and mailed off to randomly selected taxpayers. All letters are expected to be delivered to taxpayers by the latest August 12, 2016	
August 12 - 19, 2016	Send letters for sub-experiment two
Letters for sub-experiment two were generated, printed and mailed off to randomly selected taxpayers. All letters are expected to be delivered to taxpayers by the latest August 26, 2016	
September 30, 2016	Collect data to examine effects for the third quarter
Third quarter payments are due on September 15 Examine effects on the payment of tax arrears, approximately one month after receiving the treatments in sub-experiment 2.	
December 31, 2016	Collect data to examine effects for the fourth quarter
Fourth quarter payments are due on December 15 Examine effects on the payment of tax arrears, approximately 4 months after receiving the treatments in sub-experiment 2.	

Table C.2: Estimated Cost of Sub-experiment Two

Item	Unit Cost (J\$)	Total Cost (J\$)
<i>Cost of Materials</i>		
Envelopes	44.00	252,516.00
Paper	10.00	57,390.00
Toner *	30.00	172,170.00
Ink *	30.00	172,170.00
Glue *	10.00	57,390.00
Stamp	60.00	344,340.00
Total Material Cost	184.00	1,055,976.00
 <i>Cost of Labour</i>		
Labour Cost	20.77	119,200.00
Total Cost	204.77	1,175,176.00

Notes: Estimates are for sub-experiment two. Costs estimates are based on costing information provided by the tax administration. We have no data on items marked with * and the associated cost estimates are best guess upper bounds.

Table C.3: Estimated Cost and Benefit by Treatment for Sub-experiment Two

Treatment	No of Taxpayers	Marginal Cost	Marginal Revenue	Total Cost	Revenue Gains
PG_std	2883	204.77	2590	590,352.40	7,466,970.00
PG_augm	2856	204.77	2166	584,823.60	6,186,096.00
				5739	13,653,066.00

Notes: Estimates are for sub-experiment two. Table breaks out the estimated costs and revenue gains for each treatment.

Figure C.1: Envelope in which Letters were sent



Figure C.2: Sample Letter sent in Sub-experiment One, Standard Public Goods Treatment

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

February 20, 2016

Re: Taxes finance important public services provided by the Government

Dear [REDACTED]

Paying your taxes provide the funding necessary to pay for important public services such as health care, education, national security and other important services. If all self-employed individuals pay their income tax liabilities in full, this could provide funding to significantly increase spending in these important areas.

You are reminded that quarterly payments of your estimated income tax liability for 2016 are due on March 15, June 15, September 15 and December 15 of the current year. Final payment of all outstanding income tax liabilities for 2016 must be made on or before March 15, 2017. The deadline for filing your final income tax return for 2016 is March 15, 2017. This may be done on line via the TAJ website at www.jamaicatax.gov.jm or at any tax office.

We encourage you to file and pay your taxes on time and in full. You can call our Customer Care Centre at 1-888-Tax-Help (829-4357) or email us at taxhelp@taj.gov.jm if you have any questions, comments or suggestions about how we can better serve you.

Yours Sincerely,

.....
Commissioner General

Figure C.3: Sample Letter sent in Sub-experiment Two, Standard Public Goods Treatment

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

February 20, 2016

Re: Taxes finance important public services provided by the Government

Dear [REDACTED]

You are among a group of taxpayers identified as having outstanding income tax liabilities.

Paying your taxes provide the funding necessary to pay for important public services such as health care, education, national security and other important services. If all self-employed individuals pay their income tax liabilities in full, this could provide funding to significantly increase spending in these important areas.

If you have not already paid your outstanding income tax liability, please do so now.

We encourage you to file and pay your taxes on time and in full. This may be done on line via the TAJ website at www.jamaicatax.gov.jm or at any tax office. You can call our Customer Care Centre at 1-888-Tax-Help (829-4357) or email us at taxhelp@taj.gov.jm if you have any questions, comments or suggestions about how we can better serve you.

Yours Sincerely,

.....
Commissioner General

Figure C.4: Sample Letter sent in Sub-experiment Two, Augmented Public Goods Treatment

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

February 20, 2016

Re: Taxes finance important public services provided by the Government

Dear [REDACTED]

You are among a group of taxpayers identified as having outstanding income tax liabilities.

Paying your taxes provide the funding necessary to pay for important public services such as health care, education, national security and other important services. Last year the Government spent approximately J\$54 billion on health, J\$90 billion on education and J\$57 billion on national security. If all self-employed individuals pay their income tax liabilities in full, this could provide funding to significantly increase spending in these important areas.

If you have not already paid your outstanding income tax liability, please do so now.

We encourage you to file and pay your taxes on time and in full. This may be done on line via the TAJ website at www.jamaicatax.gov.jm or at any tax office. You can call our Customer Care Centre at 1-888-Tax-Help (829-4357) or email us at taxhelp@taj.gov.jm if you have any questions, comments or suggestions about how we can better serve you.

Yours Sincerely,

.....

Commissioner General

References

- Abbas, S. M., Belhocine, N., ElGanainy, A. A., & Horton, M. (2010). *A Historical Public Debt Database*. (IMF Working Paper No. 10/245). Washington DC: International Monetary Fund.
- Abiad, A., & Ostry, J. D. (2005). *Primary Surpluses and Sustainable Debt Levels in Emerging Market Countries*. (IMF Policy Discussion Paper No. 05/6). Washington DC: International Monetary Fund (IMF) .
- Afanzo, A., & Hauptmeier, S. (2009). *Fiscal Behavior in the European Union: Rules Fiscal Decentralization and Government Indebtedness*. (ECB Working Paper. No 1054). European Central Bank.
- Alberola, E., & Montero, J. M. (2006). *Debt Sustainability and Procyclical Fiscal Policies in Latin America*. (Research Paper No. WP-0611). Banco de Espana .
- Alesina, A., & Ardanga, S. (2010). *Large Changes in Fiscal Policy: Taxes versus Spending*. (NBER Working Paper No 15438). Cambridge, MA: National Bureau of Economic Research.
- Allingham, M. G., & Sandmo, A. (1972). Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, 323-338.
- Alm, J. (2012). Measuring, Explaining and Controlling Tax Evasion: Lessons from Theory, Experiments and Field Studies. *International Tax and Public Finance*, 54-77.
- Alm, J., & Martinez-Vazquez, J. (2003). Institutions, Paradigms and Tax Evasion in Developing and Transitioning Countries. In *Public Finance in Developing Countries* (pp. 147-178). Cheltenham, UK; Northampton MA, Georgia, USA: Edward Elgar. Retrieved from http://scholarworks.gsu.edu/cgi/viewcontent.cgi?article=1041&context=econ_facpub
- Alm, J., Bahl, R., & Murray, N. (1990). Tax Structure and Tax Compliance. *Review of Economics and Statistics*, 72(4), 603-613.
- Alm, J., Chetty, T., Jones, M., & McKee, M. (2011). Taxpayer Information Assistance Services and Tax Compliance Behaviour. *Journal of Economic Psychology*, 31(4), 577-586.
- Almunia, M., & Lopez-Rodriguez, D. (2014). *Heterogeneous Responses to Effective Tax Enforcement: Evidence from Spanish Firms*. (Working Paper, WP 14/12). Oxford University Centre for Business Taxation.
- Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax Compliance. *Journal of Economic Literature*, 36(2), 816-860.

- Arellano, M., & Bond, S. (1991). Some Specification Tests for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(2), 277-97.
- Arellano, M., & Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics*, 68(1), 29-51.
- Ariel, B. (2012). Deterrence and Moral Persuasion Effects on Corporate Tax Compliance: Findings from a Randomized Control Trial. *Criminology*, 50(1), 27-69.
- Auerbach, A. J. (2002). *Is There a Role for Discretionary Fiscal Policy*. (NBER Working Paper No. 9306). Cambridge, MA: National Bureau of Economic Research.
- Baldacci, E., Gupta, S., & Mulas-Granados, C. (2009). *How Effective is the Fiscal Policy Response in Systemic Banking Crises?* (IMF Working Paper No. 09/160. Washington DC: International Monetary Fund.
- Baldacci, E., Gupta, S., & Mulas-Granados, C. (2010). *Restoring Debt Sustainability after Crisis: Implications for the Fiscal Mix*. (IMF Working Paper No. 10/232). Washington DC: International Monetary Fund.
- Baldacci, E., Gupta, S., & Mulas-Granados, C. (2013). *Debt Reduction, Fiscal Adjustment and Growth in Credit Constrained Economies*. (IMF Working Paper No. 13/238). Washington DC: International Monetary Fund.
- Barro, R. J. (1979). On the Determinants of Public Debt. *Journal of Political Economy*, 87(5), 940-971.
- Becker, G. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2), 169-217.
- Besley, T., & Persson, T. (2014). Why do Developing Countries Tax so Little. *Journal of Economic Perspectives*, 34(4), 99-120.
- Bird, R., & Martinez-Vazquez, J. (2008). *Tax Effort in Developing Countries and High Income Countries: The Impact of Corruption Voice and Accountability*. Atlanta GA: Georgia State University.
- Blumenthal, M., Christian, C., & Slemrod, J. (2001). Do Normative Appeals Affect Tax Compliance? *National Tax Journal*, 54(1), 125-138.
- Blundell, R., & Bond, S. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*, 87(1), 115-143.
- Bohn, H. (2008). The Sustainability of Fiscal Policy in the United States. In R. Neck, & J. E. Sturum, *Sustainability of Public Debt and Deficits* (pp. 15-49). Cambridge MA: MIT Press.
- Bohn, H. (1998). The Behavior of US Public Debt and Deficits. *Quarterly Journal of Economics*, 113(3), 949-63.

- Bohnet, I., & Frey, B. (1999). Social Distance and Other Regarding Behavior in Dictator Games: Comment. *American Economic Review*, 89(1), 335-339.
- Bott, K., Cappelen, A. W., Sorensen, E. O., & Tungodden, B. (2014). *You've Got Mail: A Randomised Field Experiment on Tax Evasion*. (NSE Working Paper No. 26/2014). Department of Economics, Norwegian School of Economics.
- Braithwaite, V., Smart, M., & Reinhart, M. (2006). *Why Young People are Less Compliant on Tax: Enduring or Transient Defiance*. (Reg Net Occasional Paper No. 11). Regulatory Institutions Network, Australian National University.
- Brockmeyer, A., Hernandez, M., Kettle, S., & Spencer, S. (2016). *Casting the Tax Net Wider: Experimental Evidence from Costa Rica*. (World Bank Working Paper Series). Washington DC: World Bank.
- Burbidge, J. B., Magee, L., & Robb, A. L. (1988). Alternative Transformations to Handle Extreme Values of the Dependent Variable. *Journal of the American Statistical Association*, 83(401), 123-127.
- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014a). Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs. *Econometrica*, 82(6), 2295-2326.
- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014b). Robust Data-Driven Inference in the Regression Discontinuity Design. *Stata Journal*, 14(4), 909-946.
- Carillo, P., Pomeranz, D., & Singhal, M. (2014). *Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement*. (Harvard Business School Working Paper No. 15-026). Harvard Business School.
- Castro, L., & Scartascini, C. (2013). *Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment*. (IDB Working Paper IDO-WP-472). Washington DC: Inter-American Development Bank.
- Celasun, O., & Kang, J. S. (2006). *On the Properties of Various Estimators for Fiscal Reaction Functions*. (IMF Working Paper No.06/182). Washington DC: International Monetary Fund.
- Celasun, O., Debrun, X., & Ostry, J. D. (2006). *Primary Surplus Behavior and Risks to Fiscal Sustainability in Emerging Market Countries: A 'Fan Chart' Approach*. (IMF Working Paper No. 06/67). Washington DC: International Monetary Fund.
- Cryder, C., Loewenstein, G., & Scheines, R. (2013). The Donor is in the Details. *Organizational Behavior and Human Decision Processes*, 120, 15-23.
- Del Carpio, L. (2014, June 12). *Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru*. Retrieved from http://scholar.princeton.edu/sites/default/files/Arethe_neighbors_cheating_Apr2014_0.pdf.

- Devarajan, S., Swaroop, V., & Zou, H.-f. (1996). The Composition of Public Expenditure and Economic Growth. *Journal of Monetary Economics*, 37, 313-344.
- Djawadi, B. M., & Fahr, R. (2013). *The Impact of Tax Knowledge and Budget Spending Influence on Tax Compliance*. (IZA Discussion Paper No. 7255). Bonn, Germany: Institute of Labor Economics.
- Dorrenberg, P., & Schmitz, J. (2015). *Tax Compliance and Information Provision*. (IZA Discussion Paper No. 9013). Bonn, Germany: Institute of Labor Economics.
- Feld, L. P., & Frey, B. S. (2002). Trust Breeds Trust: How Taxpayers are Treated. *Economics of Governance*, 3(2), 87-99.
- Flood, R., & Marion, N. P. (2006). *Getting Shut out of the International Capital Markets: It doesn't Take Much*. (IMF Working Paper No. 06/144). Washington DC: International Monetary Fund.
- Gali, J., & Perotti, R. (2003). Fiscal Policy and Monetary Integration in Europe. *Economic Policy*, 18(37), 533-72.
- Gangl, K., Hofman, E., & Kirchler, E. (2015). Tax Authorities Interaction with Taxpayers: A Conception of Compliance in Social Dilemmas by Power and Trust. *New Ideas in Psychology*, 37, 12-25.
- Gangl, K., Torgler, B., Kirchler, E., & Hofman, E. (2015). Effects of Supervision on Tax Compliance: Evidence from a Field Experiment in Austria. *Economic Letters*, 123(3), 378-382.
- Gemmell, M., & Ratto, M. (2012). Behavioural Responses to Taxpayer Audits: Evidence from Random Taxpayer Inquiries. *National Tax Journal*, 65(1), 22-58.
- Ghosh, A., Kim, J., Mendoza, E., Ostry, J., & Qureshi, M. (2011). *Fiscal Fatigue, Fiscal Space and Debt Sustainability in Advanced Economies*. (NBER Working Paper No. 16782). Cambridge MA: National Bureau of Economic Research.
- Gupta, A., & Gupta, A. S. (2007). *Determinants of Tax Revenue Efforts in Developing Countries*. (IMF Working Paper No. 07/184). Washington DC: International Monetary Fund.
- Hallsworth, M., List, J., Metcalfe, R., & Valev, I. (2014). *The Behaviouralist as a Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance*. (NBER Working Paper 20007). Cambridge MA: National Bureau of Economic Research.
- Hansen, B. E. (1999). Threshold Effects in Non-Dynamic Panels: Estimation, Testing and Inference. *Journal of Econometrics*, 93(2), 345-368.
- Hansen, B. E. (2000). Sample Splitting and Threshold Estimation. *Econometrica*, 68(3), 575-603.

- Imbens, G. W., & Kalyanaraman, K. (2009). *Optimal Bandwidth Choice for the Regression Discontinuity Estimator*. (NBER Working Paper No. 14726). Cambridge MA: National Bureau of Economic Research.
- Keen, M. (2012). *Taxation and Development - Again*. (IMF Working Paper 12/220). Washington DC: International Monetary Fund.
- Kirchler, E., & Wahl, I. (2010). Tax Compliance Inventory: Tax-i Voluntary Tax Compliance, Enforced Tax Compliance, Tax Avoidance and Tax Evasion. *Journal of Economic Psychology*, 31(3), 331-346.
- Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus Voluntary Tax Compliance: The 'Slippery Slope' Framework. *Journal of Economic Psychology*, 29(2), 210-225.
- Klemm, A. (2014). *Fiscal Policy in Latin America over the Cycle*. (IMF Working Paper 07/184). Washington DC: International Monetary Fund.
- Kleven, H. J., & Waseem, M. B. (2012). *Behavioral Responses to Notches: Evidence from Pakistan Tax Records*. (LSE Working Paper). London School of Economics.
- Kleven, H., Knudsen, M., Kreiner, C., Pedersen, S., & Saez, E. (2011). Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica*, 79(3), 651-692.
- Kosonen, T., & Ropponen, O. (2015). The Role of Information in Tax Compliance: Evidence from a Natural Field Experiment. *Economics Letters*, 129, 18-21.
- Krogstrup, S. (2002). *Public Debt Asymmetries: The Effect on Taxes and Spending in the European Union*. (ECB Working Paper No. 162). European Central Bank.
- Lamberton, C., De Neve, J. E., & Norton, M. (2014). *Eliciting Taxpayer Preferences Increases Tax Compliance*. (CEP Discussion Paper No 1270). London: Centre for Economic Performance, London School of Economics and Political Science.
- Lane, P. R. (2003). Cyclical Behavior of Fiscal Policy: Evidence from the OECD. *Journal of Public Economics*, 87(12), 2661-2675.
- Lee, D., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2), 281-355.
- Lukkezen, J., & Roja-Romagosa, H. (2012). *When is Debt Sustainable?* (CPB Discussion Paper No. 212). CPB Netherlands Bureau of Economic Policy Analysis.
- Luttmer, E., & Singhal, M. (2014). Tax Morale. *Journal of Economic Perspectives*, 28(4), 149-168.
- Mauro, P., Romeu, R., Binder, A., & Zaman, A. (2013). *A Modern History of Fiscal Prudence and Profiligacy*. (IMF Working Paper No. 13/05). Washington DC: International Monetary Fund.

- Mckee, M., Siladke, C., & Vossler, C. (2011). *Behavioral Dynamics of Tax Compliance Under an Information Service Initiative*. (MPRA Working Paper No. 38865). MPRA.
- Mendoza, E. G., & Ostry, J. D. (2008). International Evidence on Fiscal Solvency. *Journal of Monetary Economics*, 55(6), 1081-1093.
- Ortega, D., & Sanguinetti, P. (2013). *Deterrence and Reciprocity Effects on Tax Compliance: Experimental Evidence from Venezuela*. (CAF Working Paper No. 2013/08). Development Bank of Latin America.
- Ortega, D., & Scartascini, C. (2015). *Don't Blame the Messenger: A Field Experiment on Delivery Methods for Increasing Tax Compliance*. (IDB Working Paper No. 627). Washington DC: Inter-American Development Bank.
- Paetzold, J., & Winner, H. (2014). *Taking the High Road? Compliance with Commuter Tax Allowances and the Role of Evasion Spillovers*. (Oxford Center for Business Taxation Working Paper No. 14/19). Oxford Center for Business Taxation.
- Peralta, A., Tavares, M. M., & Tam, X. S. (2016). *The Distributional Implications of Fiscal Consolidation in Developing Countries*. Retrieved December 10, 2016, from https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=CSAE2016&paper_id=655
- Perez-Truglia, R., & Troiano, U. (2015). *Shaming Tax Delinquents: Theory and Evidence from a Field Experiment in the United States*. (NBER Working Paper No. 21264). Cambridge MA: National Bureau of Economic Research.
- Plödt, M., & Reicher, C. (2014). *Estimating Simply Fiscal Reaction Functions for the EURO Area Countries*. (Kiel Working Paper No. 1899). Kiel Institute for the World Economy.
- Pomeranz, D. (2013). *No Taxation Without Information: Deterrence and Self-enforcement in the Value Added Tax*. (NBER Working Paper No. 19199). Cambridge MA: National Bureau of Economic Research.
- Reicher, C. P. (2013). *A Set of Estimated Fiscal Rules for a Cross Section of Countries: Stabilization and Consolidation through Which Instruments?* (Kiel Working Paper No. 1850). Kiel Institute for the World Economy.
- Reinhart, C., & Sbrancia, M. (2011). *The Liquidation of Government Debt*. (NBER Working Paper No. 16893). Cambridge MA: National Bureau of Economic Research.
- Reinhart, C., Rogoff, K., & Savastano, M. (2003). *Debt Intolerance*. (NBER Working Paper No. 9908). Cambridge MA: National Bureau of Economic Research.
- Saez, E. (2010). Do Taxpayers Bunch at Kink Points. *American Economic Journal: Economic Policy*, 2(3), 180-212.
- Sanchez, G. E. (2013). *The Impact of Low Cost Intervention on Tax Compliance: Regression Discontinuity Evidence*. (Unpublished Manuscript). University of California, Berkeley.

- Slemrod, J., & Blumenthal, M. (2001). Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota. *Journal of Public Economics*, 79(3), 455-483.
- Tax Administration Jamaica. (2015). *Tax Administration of Jamaica National Compliance Plan*. Kingston, Jamaica: Tax Administration Jamaica (TAJ). Retrieved from https://www.jamaicatax.gov.jm/documents/10181/786863/taj_ncp_2015-2016_presentation.pdf
- Torgler, B. (2004). Moral Suasion: An Alternative Tax Policy Strategy? Evidence from a Controlled Field Experiment in Switzerland. *Economics of Governance*, 5(3), 235-253.
- Vossler, C. A., & Mckee, M. (2013). *Efficient Tax Reporting: The Effect of Taxpayer Information Services*. (Appalachian State Department of Economics Working Paper No. 13-24). Appalachian State University, Department of Economics.
- Wenzel, M. (2005). Misperception of Social Norms About Compliance: From Theory to Intervention. *Journal of Economic Psychology*, 26(6), 862-883.

VITA

Oronde is a native of Jamaica and holds a B.Sc. and M.Sc. degree in economics from the University of the West Indies, Mona (UWI). In 2004, after completing his M.Sc., Oronde worked at the Ministry of Finance in Jamaica as an economist in the Research and Analysis Unit of the Tax Policy Division. In 2008 he assumed the role of Director, International Trade Relations in the same Division, and served in this position for 4 years, before leaving to further his studies.

Oronde entered the economics doctoral program at the Andrew Young School of Policy Studies (AYSPS), Georgia State University in 2012. During his time at AYSPPS Oronde worked as a graduate research assistant in the Fiscal Research Center (FRC). The experience and exposure gained from working alongside renowned public finance experts at FRC has helped to shape his academic and policy related research interests. Oronde's primary research interests are in the fields of public finance, health and experimental economics. He has presented his research at several conferences including some hosted by the National Tax Association, Southern Economic Association and the University of the West Indies.

Oronde will graduate with a Doctor of Philosophy degree in Economics from the Andrew Young School of Policy Studies, Georgia State University in August 2017 and will resume his duties at the Ministry of Finance in Jamaica.