The elasticity of taxable income: the case of South Africa

Johannes Hermanus Kemp
Bureau for Economic Research, Stellenbosch University

May 18, 2017

Abstract

A central tax policy parameter that has received much attention internationally, but about which there is substantial uncertainty, is the overall elasticity of taxable income. The size of this elasticity is of critical importance in the formulation of tax and transfer policy, as well as for the study of the welfare implications of tax decisions. This paper uses a panel of individual tax returns for the period 2009 - 2013 and the phenomenon of 'bracket creep' as source of tax rate variation to construct instrumental variable estimates of the sensitivity of income to changes in tax rates. We find that the overall elasticity of taxable income is approximately 0.3, while that of broad income is significantly lower. This overall elasticity is primarily due to the elastic response of taxable income for taxpayers who have incomes above R380 000, who have an elasticity of closer to 0.4. The estimates suggest an optimal marginal tax rate for the top 10% of income earners that is generally in line with the current income tax schedule. However, results also suggest that there is little scope for raising marginal rates on high-income earners without inducing a negative revenue response.

JEL Classification: H21; H31; J22

Keywords: fiscal policy; elasticity; taxable income; optimal tax

1 Introduction

The response of taxpayers to changes in marginal tax rates has long been of interest to economists. The magnitude of this response is of critical importance in the formulation of tax and transfer policy, as well as for the study of the welfare implications of tax decisions. Public finance practitioners are often asked to predict responses to alternative policy changes and to provide empirical estimates of the effects to decision-makers (Thoresen and Vattø, 2013). **To this end, it is important to quantify the behavioural response of taxpayers to changing policy parameters.**

In this regard, the elasticity of taxable income, or ETI, is a key concept. The ETI aims to capture all possible behavioural responses to changes in income taxation, including tax evasion and/or avoidance, in a single measure, without the need to specify the nature of the specific adjustment processes involved (Creedy, 2009; Thoresen and Vattø, 2013).[1] **This allows policymakers to evaluate the likely impact of policy proposals on both taxpayer behaviour and policy outcomes (such as the impact on tax revenue of a change in marginal tax rates).**

Modelling the response to tax rates in this reduced form way, as opposed to formulating a structural model of the behaviour involved, has proven to be very attractive. Since the early 1990's, a large literature has sought to estimate the ETI and elasticities for related income measures. Most of this literature has focused on individual taxable income in the United States, although some studies have examined taxable income elasticities in Canada and Western Europe, while a limited number of studies have focused on countries outside these regions.

Very little work has been done in the South African context. Given the lack of access to micro-level panel data on personal income taxes (PIT), studies that have attempted to estimate the ETI for South Africa focus primarily on publicly available aggregated tax data. Using a new dataset comprising confidential tax return data made available for research purposes by the South African Revenue Service (SARS) and the National Treasury, this paper attempts to estimate the ETI for South Africa.

Much of the literature ETI looks directly at the sensitivity of overall income with respect to marginal tax rates using tax reforms to identify the parameters of interest.[2] Notwithstanding the fact that the ETI varies between countries depending on the nature of the tax system, the empirical literature has failed to generate a consensus of the magnitude of the elasticity even within specific countries.

According to Saez (2003), two reasons might explain this lack of consensus. First, most tax reforms introduce numerous changes to the definition of taxable income besides tax rate changes. This makes it difficult to compare reported income before and after a tax reform. Second, most studies rely on a comparison between high- and low- or middle-income taxpayers. The former often experience large tax rate changes after any given tax reform, while the latter experience little or no change in marginal tax rates.

This suggests that the research design for estimating behavioural responses to changes in marginal tax rates should meet two conditions. First, the tax change should affect only marginal tax rates without introducing changes in tax rules. Second, the tax change should affect differently groups of taxpayers that are comparable in terms of incomes and other economic characteristics.

Taking these factors into account and considering the fact that there was no large legislated tax reforms in South Africa over our sample period, we follow Saez (2003) in utilising the phenomenon of 'bracket creep' to estimate the behavioural response of taxpayers with respect to changing marginal tax rates. Because of inflation, a taxpayer near the top end of an income tax bracket might 'creep' to the next bracket even if their real income does not change. This is especially true if nominal

[1]These adjustments could include labour supply changes (at both the extensive and/or intensive margins), income shifting between sources which are taxed at different rates, and tax evasion through non-declaration of income.

[2]It should be noted that while the larger share of the literature has focused on personal income taxes (PIT) and the response of individual taxable income to changes in marginal tax rates, most of the conceptual and empirical issues discussed remain relevant for other tax bases, such as corporate income taxes (CIT).

income tax brackets are not fully adjusted for inflation (so-called fiscal drag). In contrast, taxpayers at the bottom end of the relevant tax bracket are not as likely to experience an increase in marginal rates.

In practice, the method compares changes in the incomes of taxpayers near the top end of a bracket to those taxpayers at the bottom end of the bracket using tax return data. Importantly, 'bracket creep' does not affect the definition of reported income. Therefore incomes can be easily compared across years. Additionally, this strategy compares groups of taxpayers that are closely related in terms of income levels and economic characteristics.

Two limitations of the approach should be mentioned. First, changes in tax rates due to 'bracket creep' were relatively small compared to other tax reforms across the decades. Second, because 'bracket creep' is not a legislated tax change, it might be harder for taxpayers to understand the effect of this change on marginal tax rates. These two caveats suggest that the ETI estimates using 'bracket creep' reflects a lower bound for the behavioural response to changing tax rates. We will discuss and dissect these important limitations throughout the paper. Nevertheless, in the absence of any large legislated tax reforms, 'bracket creep' provides a useful natural experiment to look for direct evidence of behavioural responses to taxation.

The rest of the paper proceeds as follow. Section 2 discusses the basic conceptual framework and unpacks the various identification issues that arise when attempting to estimate the ETI in practice. Section 2 also provides a short review of the international empirical literature. Section 3 estimates the ETI for South Africa using micro-level tax return data and Section 4 discusses optimal taxation results. Section 5 presents some caveats and concludes.

2 Conceptual Framework and Literature Review

Modelling the behavioural response to income taxation has taken on two distinct forms. The first is based on the standard labour supply model and attempts to model behavioural responses in a structural way. Using cross-sectional observations on households' and individuals' consumption and connections to the labour market (typically through a measure of hours worked), it is possible to directly apply a labour supply function or estimate a utility function.[3] These parameter estimates can, in turn, be used to simulate the effects on hours worked and/or income of hypothetical changes to the tax system (Thoresen and Vattø, 2013). This approach forms the basis for much of the literature on microsimulation tax models.

While such discrete-choice labour supply models remain key instruments in predicting the effects of policy changes, concerns have been raised about their ability to generate robust predictions (see LaLonde (1986), Angrist and Pischke (2010) and Imbens (2010), among others). **A major concern is the stability of estimated structural parameters in the face of major policy shifts. Keane and Wolpin (2007) argue that there is no true decision-theoretic model, only models that are better or worse at addressing particular questions.** As such, it is essential to use other sources of information to validate the models. For example, Blundell (2006) notes that "...simple difference in difference evaluations can be valuable for validating the specification of more fragile microeconometric models" (p. 425). **In fact, a large literature has developed around the validation and estimation of so-called structural microeconomic models using quasi-experimental designs (see, for example, Brewer et al. (2006), Hansen and Liu (2011) and Thoresen and Vattø (2013)). Given the fact that there is little agreement in the labour supply literature on how responsive individuals are to changes in taxes, even in the static case (see, for example, Keane and Rogerson (2012)), reconciliation of evidence from the different approaches is valuable.**

[3]See Blundell and Macurdy (1999) for a review of the literature

This brings us to the second broad approach to modelling the behavioural response to changing taxation, as pioneered by Lindsey (1987) and Feldstein (1995). It is mainly based on the analysis of observations before and after a realised policy reform. In this tradition, policy reforms are seen as quasi-experiments. A tax reform generates changes in net-of-tax incomes along the entire income distribution, often resulting substantial tax changes for some taxpayers while leaving others largely unaffected. As such, the identification of response estimates typically involves the application of the difference-in-differences estimator or other related techniques (such as the instrumental variables, or IV, approach).

The approach has been extended to incorporate identification strategies not necessarily focused on specific tax reforms, including using the phenomenon of 'bracket creep' to obtain elasticity estimates (Saez, 2003) and the use of bunching analysis to estimate income responses around kink points in the tax schedule (developed by Saez (2010) and extended by Chetty et al. (2011), among others).

While the reduced form approach has gained in popularity, it has limited value in a prediction context because the estimation relies on a particular reform for identification, making it less useful when assessing new policy changes.[4] Importantly, the elasticity for a given individual or group may not be constant and depends on the tax system. As a result, an elasticity estimated around the current tax system may not apply to a hypothetical large tax change (Saez et al., 2012).

Despite this apparent shortcoming, the relative simplicity of the approach has proven to be attractive. It uses information about incomes, which is normally more easily accessible than data on hours worked, and exploits standard econometric techniques. Additionally, it has been suggested that the empirical method in the ETI literature can be used to identify earnings responses that can be compared to those obtained from microsimulations based on the structural labour supply models, which can then serve as a form of validation for the latter (Thoresen and Vattø, 2013).

We will now turn to a brief description of the basic model employed in the reported income, or ETI, literature, followed by a discussion of the various identification issues that crop up in empirical exercises. A brief literature overview will highlight the numerous ways in which practitioners have attempted to solve these identification problems.

2.1 Basic Model[5]

In the standard labour supply model, individuals maximise a utility function $u(c, l)$, where c is consumption (equal to disposable income in a one-period model), and l is labour supply (measured by hours worked). Earnings are given by $w \cdot l$, where w is the exogenous wage rate. The (linearised) budget constraint is $c = w \cdot l \cdot (1 - \tau) + R$, where τ is the marginal tax rate and R is virtual income (that is, non-wage income equal to the intercept of the linear section of the budget constraint extended to the c axis, where $l = 0$).

The ETI literature generalises this model by noting that hours worked is just one of the possible behavioural responses to income taxation. Others include intensity of work, career choice, form and timing of compensation, and tax avoidance, or even tax evasion. As a result, an individual's wage rate w might depend on effort and also respond to tax rates. Reported taxable income could therefore differ from $w \cdot l$ as individuals might split their gross income between taxable (cash) compensation and non-taxable compensation such as fringe benefits, or even fail to report their full taxable income because of tax evasion (Saez et al., 2012).

As shown in Feldstein (1999), a simple way to jointly model these responses is to assume that utility depends positively on disposable income (equal to consumption) c and negatively on reported or taxable income z (activities that generate income are costly because they might, for example, require

[4] There are exceptions. See Carroll and Hrung (2005) and Thoresen et al. (2012), among others.

[5] **Following Saez (2001), Gruber and Saez (2002) and Saez et al. (2012)**

foregoing leisure).[6] As such, individuals maximise utility $u(c,z)$ subject to the linearised budget constraint $c = z\cdot(1-\tau)+R$, where z is before tax (taxable) income and R is virtual income. Utility maximisation leads to an individual 'reported income' supply function which depends on the slope of the budget line and virtual income: $z = z(1-\tau, R)$. Changes in τ and R affect reported income as follows:

$$dz = -\frac{\partial z}{\partial(1-\tau)}d\tau + \frac{\partial z}{\partial R}dR \tag{1}$$

Following Gruber and Saez (2002) and introducing the uncompensated elasticity of reported (taxable) income with respect to the net-of-tax rate $e^u = \frac{(1-\tau)}{z}\frac{\partial z}{\partial(1-\tau)}$ and the income effect parameter $\eta = \frac{(1-\tau)\partial z}{\partial R}$, we have:

$$dz = -e^u z\frac{\partial \tau}{1-\tau} + \eta\frac{dR}{1-\tau} \tag{2}$$

Introducing the compensated ETI, e^c, and using the Slutsky equation $e^c = e^u - \eta$, we obtain finally:

$$\frac{dz}{z} = -e^c\frac{d\tau}{1-\tau} + \eta\frac{dR - zd\tau}{z(1-\tau)} \tag{3}$$

where $dR - zd\tau$ is the change in after-tax income due to the tax change for a given before-tax income z, or the change in tax liability for taxpayers with income z. Equation (3) therefore decomposes the change in reported income in response to a change in the marginal tax rate into a substitution effect (or the behavioural response of the tax-payer) and an income effect.

In much of the ETI literature, income effects are assumed away so that the income supply function $z(\cdot)$ does not depend on virtual income R, but only on the net-of-tax rate, $1-\tau$. This equates to setting $\eta = 0$ in (3). As mentioned in Saez et al. (2012), there is no consensus in the labour supply literature on the size of income effects, with many studies finding small and insignificant effects, but some important studies finding large income effects (see Blundell and Macurdy (1999) for a survey). In the ETI literature, there is much less empirical evidence on the size of income effects.[7]

If one assumes away income effects (i.e. $e^c = e^u = e$), the goal of the ETI literature amounts to estimating the elasticity of taxable (or reported) income with respect to the net-of-tax rate, defined as

$$e = \frac{(1-\tau)}{z}\frac{\partial z}{\partial(1-\tau)} \tag{4}$$

or the percent change in reported income when the net-of-tax rate increases by 1 percent. As shown in Feldstein (1999), this elasticity captures not only the hours of work response, but also all other responses to changes in marginal tax rates.

Importantly, the elasticity is not a structural parameter that depends solely on individual preferences. It depends crucially on the features of the tax system, such as opportunities for deductions of other avoidance mechanisms. Therefore, the elasticity for a particular individual may not be constant over time and depends on the structure and nature of the tax system.

Several issues arise when attempting to estimate the ETI in practice. We now turn to a discussion on some of the identification issues that arise in the empirical literature.

[6] **Note that it is not necessary to define a specific functional form for the utility function. It is sufficient to require that it is increasing in consumption (disposable income) and decreasing in reported income.**

[7] Gruber and Saez (2002) estimate both income and substitution effects in the case of reported taxable income for the US and find small and insignificant income effects.

2.2 The Main Identification Problem

A simple model of income reporting behaviour can serve to illustrate some of the identification issues in the estimation of the ETI. Following Saez et al. (2012), it is assumed that in year t, individual i reports income z_{it} and faces a marginal tax rate $\tau_{it} = T'(z_{it})$. Reported income is assumed to respond to marginal tax rates with elasticity e so that $z_{it} = z_{it}^0 \cdot (1 - \tau_{it})^e$, where z_{it}^0 is income reported when the marginal tax rate is zero, i.e. potential income. Using logs we have:

$$\log z_{it} = e \cdot \log(1 - \tau_{it}) + \log z_{it}^0 \tag{5}$$

Note that there are numerous assumptions built into this simple model. Firstly, there are no income effects on reported income. Second, the response to tax rates is immediate and permanent (i.e. the short-term and long-term elasticities are identical). Thirdly, the elasticity e is constant over time and uniform across individuals at all levels of income. Finally, individuals have perfect knowledge of the tax system and choose z_{it} after they know the exact realisation of z_{it}^0.[8]

Nevertheless, this simple framework serves to illustrate the main identification issue highlighted in the literature. Even in the context of this simple model, an OLS regression of $\log z_{it}$ on $\log(1 - \tau_{it})$ would not identify the elasticity e in the presence of a graduated income tax schedule. This is because τ_{it} is positively correlated with $\log z_{it}^0$ because the marginal rate might increase with an increase in realised income. Therefore, it is necessary to find instruments correlated with τ_{it}, but uncorrelated with $\log z_{it}^0$ to identify the elasticity e.

Much of the empirical literature has used changes in the tax rate structure created by tax reforms to obtain such instruments. By comparing the observed reported incomes after the tax rate change to the incomes that would have been reported had the tax change not taken place, one can isolate the effects of the change in the net-of-tax rate (Saez et al., 2012). Obviously, these counterfactual incomes are not observed and need to be estimated - an issue that has received much attention in the ETI literature.

One simple approach uses reported incomes before a tax reform as a proxy for reported incomes after the reform. Essentially, this approach boils down to comparing reported incomes before and after the reform and attributing the change in reported incomes to the changes in tax rates. In many cases, however, a tax reform affects different subgroups of the population differently, and in some cases they leave tax rates essentially unchanged for most of the population. In this context, the group less affected by the reform can be used as a control group and can therefore be used to proxy the unobserved income changes for the affected group. Such methods naturally lead to the consideration of difference-in-differences estimation. In fact, most of the recent literature on ETI has used micro-based regressions using these difference-in-differences methods (Saez et al., 2012).

We will look at two widely used techniques in the ETI literature.

2.2.1 Repeated Cross-Section Analysis

In this approach, changes in reported incomes for a treatment group, T, (i.e. a group affected by the tax change) are compared to changes for a control group, C, not affected by the tax change. Denote by t_0 and t_1 the pre- and post-reform years respectively. Then, the following two-stage-least squares regression can be estimated:

$$\log z_{it} = e \cdot \log(1 - \tau_{it}) + \alpha \cdot I(t = t_1) + \beta \cdot I(i \in T) + \varepsilon_{it} \tag{6}$$

where $I(\cdot)$ is the indicator function which takes the value of 1 if the condition in brackets is met. The regression in (6) is estimated on a repeated cross-section sample including both the treatment

[8] These assumptions can be relaxed in most cases, but it sometimes has important implications for identification.

and control groups and including the year t_0 and year t_1 samples, and using as an instrument for $\log(1-\tau_{it})$ the post-reform and treatment group interaction $I(t=t_1)\cdot I(i\in T)$ (Saez et al., 2012). The two-stage-least squares estimate is a classical difference-in-differences estimator equal to:

$$e = \frac{([E(\log z_{it_1}|T) - E(\log z_{it_0}|T)] - [E(\log z_{it_1}|C) - E(\log z_{it_0}|C)])}{([E(\log(1-\tau_{it_1})|T) - E(\log(1-\tau_{it_0})|T)] - [E(\log(1-\tau_{it_1})|C) - E(\log(1-\tau_{it_0})|C)])} \tag{7}$$

Therefore, the elasticity estimate is defined as the ratio of pre- to post-reform change in log incomes in the treatment group minus the same difference for the control group, to the same difference-in-differences in the log net-of-tax rates. As mentioned in Saez et al. (2012), the elasticity estimate will be unbiased only if the parallel trend assumption holds: absent the tax change, the numerator would have been zero, i.e. log-income changes pre- to post-reform would have been the same for both the control and treatment group.

This brings us to a general concern in the ETI literature regarding the identification of elasticity estimates, namely the contamination of ETI estimates through secular changes in the (taxable) income distribution unrelated to changes in the tax structure. Reported incomes (for both treatment and control groups) can change over time for reasons unrelated to tax changes, which might lead to biased estimates for the elasticity of taxable income. One way to control for changes in the income distribution is to include separate time trends for the treatment and control groups in the regression framework.

2.2.2 Panel Analysis

Following the influential analysis of Feldstein (1995), the great majority of empirical studies of the ETI have used panel data. With panel data, we can follow individual tax filers between pre- and post-reform years in order to identify income responses to changes in tax rates.[9] Estimation of the ETI is carried out by running the two-stage-least-squares panel regression:

$$\log\left(\frac{z_{it_1}}{z_{it_0}}\right) = e\cdot\log\left(\frac{1-\tau_{it_1}}{1-\tau_{it_0}}\right) + \varepsilon_{it} \tag{8}$$

using the treatment indicator $I(i\in T)$ as an instrument for the log difference in the net-of-tax rates.[10] This regression estimates a difference-in-differences parameter:

$$e = \frac{E\big(\log\big(\frac{z_{it_1}}{z_{it_0}}\big)|T\big) - E\big(\log\big(\frac{z_{it_1}}{z_{it_0}}\big)|C\big)}{E\big(\log\big(\frac{1-\tau_{it_1}}{1-\tau_{it_0}}\big)|T\big) - E\big(\log\big(\frac{1-\tau_{it_1}}{1-\tau_{it_0}}\big)|C\big)} \tag{9}$$

As noted above, this elasticity estimate is unbiased if, absent the tax change, the numerator is zero. Again, this assumption might be violated if there are secular changes in the income distribution. However, with panel analysis, another problem arises. Even in the absence of changes in the income distribution, the identification assumption might be violated because of the presence of mean reversion in reported income. High incomes in year t_0 tend to be lower in the following years, especially if a large share of year t_0 income comprises positive transitory incomes. This produces a negative correlation between ε_{it} and z_{it_0} at the top of the income distribution. Similarly, mean reversion could produce a positive correlation at the bottom of the distribution.

[9] This is the strategy that will be followed in estimating the ETI for South Africa in Section 3

[10] Alternatively, as in Gruber and Saez (2002), it is possible to use an instrument equal to the predicted log change in the net-of-tax rate, $\log[(1-\tau_p)/(1-\tau_{t_0})]$, where τ_p is the predicted marginal tax rate that the individual would face post-reform if his real income did not change between pre- and post-reform years. More detail will follow when estimating the ETI for South Africa.

To mitigate the mean reversion bias as well as to control for potential changes in the income distribution, the recent literature has followed Auten and Carroll (1999) in introducing t_0 income controls in (8), either in a simple way by including $\log z_{it_0}$ or in a richer way by including polynomials or splines in base-year incomes. However, with a small time dimension (in our example we only use two years: pre- and post-reform), including income controls could destroy identification. This is because the inclusion of income controls absorbs much of the independent variation in tax rates, which are correlated with income.

To overcome this issue, following Gruber and Saez (2002), additional observations can be added to the regression by stacking differences across multiple years and adding year dummies as follows:

$$\log\left(\frac{z_{it+1}}{z_{it}}\right) = e \cdot \log\left(\frac{1-\tau_{it+1}}{1-\tau_{it}}\right) + f(z_{it}) + \alpha_t + \varepsilon_{it}, \quad t = 0, 1, 2... \tag{10}$$

where $f(z_{it})$ denotes controls in base-year income and α_t are year-specific dummies.

Panel analysis is particularly useful in cases where the aim is to analyse a tax change targeted at a specific group and there is a concern that the composition of this group might change over time in repeated cross sections. For example, suppose a tax change is targeted at the top end of the income distribution and reported incomes of the top 1% change for unrelated reasons (i.e. new tax filers enter the top of the distribution for reasons unrelated to the tax change). In this scenario, repeated cross-section analysis is biased because of composition effects, but panel analysis is not (or at least is less so) since the same tax filers are followed over time. Additionally, panel data are required to study questions other than the overall response of reported incomes. For example, in order to study how tax changes affects income mobility, panel data is clearly necessary.

One important shortcoming of panel analysis (relative to repeated cross-section analysis) is that the identification restrictions lack transparency (Saez et al., 2012). The identification strategy discussed above mixes assumptions regarding mean reversion and assumptions regarding possible changes in the income distribution. It is often impossible to tell to what extent income controls are adjusting for mean reversion and to what extent they are controlling for divergence in the income distribution. Additionally, these income controls could hamper identification by absorbing informative variation in tax rates. To overcome this, Saez et al. (2012) suggest that it might be useful to include episodes of both increases and decreases in tax rates for the purpose of identification as mean reversion creates biases in opposite directions in the case of tax increases versus tax decreases.

Finally, it has been shown that panel regressions are relatively sensitive to the choice of instrument, suggesting that the standard methods do not control adequately for mean reversion.

2.3 Other Identification Issues

2.3.1 Alternative Control Groups

The tax code itself might offer opportunities to identify alternative control groups that may be more comparable to the treatment group. For example, inflation or the loss of a personal exemption might push tax filers into higher tax brackets, and in some cases (particularly in the US) married and unmarried couples might experience different changes in tax schedules (Saez et al., 2012).[11]

The main advantage of this approach is the ability to compare taxpayers who are very similar in terms of income and initial marginal tax rates, but face different prospects for changes in marginal rates. This makes a much more compelling case for identification.

The main drawback is that taxpayers might not be aware of the minute details of the tax code and the possible (localised) changes to their marginal tax rates. As a result, the estimated elasticities

[11] See Saez (2003), Looney and Singhal (2006), Feldman and Katuščák (2009) and Singleton (2011), among others, for examples of the use of alternative control groups.

might not be relevant for predicting behavioural responses in the case of well-publicised, broader tax changes.

This provides for the fundamental tension in tax response analysis: large tax changes are likely to be noticed and acted upon by taxpayers, but often do not generate convincing control groups for identification purposes; small changes or specific aspects of the tax code might generate better control groups, but are unlikely to produce generalisable elasticity estimates.

2.3.2 Tax Base Changes

The estimation of ETI focuses on the behavioural response of taxpayers to changes in the tax rate schedule. However, the definition of the tax base that is subjected to the rates schedule also changes periodically, often in conjunction with the change in the tax schedule. This raises a number of issues.

First, the estimated elasticity will be plausibly different in the post-reform era relative to the pre-reform era. For example, in the US, the Tax Reform Act of 1986 broadened the tax base by reducing deductions and the attractiveness of tax shelters, which resulted in a lower elasticity post-1986 than in the pre-reform era (as documented in Kopczuk, 2005). In cases such as this, even well-specified estimation strategies will yield an estimate of neither the pre- nor post-reform ETI, but rather a weighted average of the two (Saez et al., 2012).

Second, when the definition of the tax base changes, using the concurrent definition of taxable income as dependent variable in the regressions above runs the risk of confounding tax-induced behavioural changes with purely definitional changes. The best course of action is to use a consistent pre- or post-reform definition of taxable income. However, as discussed in (Saez et al., 2012), even if a consistent definition is used, the choice of which specific constant-law definition to use is not innocuous.

2.3.3 Capital Gains and Multiple Tax Rates

An additional complicating factor is capital gains. Many studies of the ETI exclude capital gains from the definition of taxable income on the grounds that these gains are more often than not subject to a different marginal tax rate than other income. Additionally, it has been found that capital gains realisations are particularly sensitive to anticipated changes in the applicable tax rate (see Burman et al., 1994). Some taxpayers have opportunities for converting wage and salary income into capital gains income. These opportunities are exercised when the tax advantages of doing so are sufficient. The advantage of doing so depends on the differential between the tax rates on ordinary income and capital gains. Additionally, converting ordinary income to capital gains income is often accompanied by a deferral of the realisation of said capital gains income for tax purposes, so the offsetting tax revenue might not materialise for several years, if ever.

As such, the capital gains rate matters and should ideally be included in the analysis of the aggregate behavioural response to changes in the tax rate on ordinary income. Ideally, a model of joint determination of capital gains and ordinary income as a function of the different marginal tax rates should be embedded in the general methodology of the elasticity of taxable income. Given the difficulties in specifying such as model, most studies exclude capital gains income from the measure of taxable (or more broadly defined) income.

2.4 Empirical Analysis

Since the early 1990s, a large literature has sought to estimate the ETI and elasticities for related income measures. Most of this literature has focused on individual taxable income in the United States, although some studies have examined taxable income elasticities in Canada and Western Europe, while a limited number of studies have focused on countries outside these regions.

We turn now to a selective review of the international literature on the estimation of the ETI.

2.4.1 ETI in the United States

Feenberg and Poterba (1993) were the first to use aggregate tax return data to shed light on the high-income share of reported income in the United States and how this might be influenced by the tax rate structure. They calculate for 1951 to 1990 the share of Adjusted Gross Income (AGI) accruing to the top 0.5% of taxpayers and find that the time series pattern is consistent with a behavioural response to reductions in the top marginal tax rate, particularly following the Tax Reform Act of 1986.

Following on the Feenberg and Poterba (1993) study, Feldstein (1995) pioneered the use of panel data in utilising the United States public-use panel tax data to estimate the ETI for the Tax Reform Act of 1986. Feldstein (1995) uses a tabulated difference-in-differences methodology and generates large ETI estimates, ranging from 1 to 3 in alternative specifications. Using Feldstein's (1995) approach, Moffitt and Wilhelm (2000) find elasticities of AGI around the 1986 tax reform ranging from 1.76 to 1.99. The results apparently corroborated the findings of Feenberg and Poterba (1993) that something extraordinary happened to the dispersion of taxable income at around the same time as the tax reform of 1986.

Subsequent research has focused mainly on the use of panel data in the tradition of Feldstein (1995), but has attempted to address some of the econometric issues raised above, including mean reversion and/or the existence of secular income trends.

Auten and Carroll (1999) address mean reversion and attempt to control for the divergence in the income distribution by including base year income controls, as well as controls for region and occupation. The authors adopt a two-stage least squares regression approach and report an elasticity estimate of 0.55. Moffitt and Wilhelm (2000) also turn to a two-stage least squares regression approach in the vein of Auten and Carroll (1999) and find tax elasticity estimates that range from 0.35 to 0.97 using alternative instruments.

However, the included controls likely failed to adequately control for mean reversion. Both studies only use two years (i.e. pre- and post-reform) and as such the results are particularly sensitive to mean reversion. The 1986 tax reform reduced top marginal tax rates in the US which likely resulted in a downward bias in estimated elasticities in the presence of mean reversion. In fact, when including 1983 AGI as a further control, Moffitt and Wilhelm (2000) find that their elasticity estimates increase by between 0.3 and 0.5.

Gruber and Saez (2002) address some of the shortcomings in earlier studies by analysing behavioural changes over three-year intervals around both the 1981 and 1986 tax reforms. The authors devote much attention to the issues of mean reversion and secular income trends and include several alternative specifications of base-year income controls. **For example, they include richer income controls in the form of splines in base year income to control for possible non-linearities in the widening of the income distribution.** At 0.4, their estimated elasticity lies in the middle of the post-Feldstein range.

Using similar methods to Gruber and Saez (2002), Auten et al. (2008) investigate the Bush tax cuts of the early 2000s. The model includes a number of additional variables in an attempt to control for key taxpayer characteristics and mean reversion. Auten et al. (2008) estimate a taxable income elasticity in the base model of about 0.4, with estimates for other specifications and samples ranging from about 0.2 to 0.7. More recently, Singleton (2011) applies similar methods to the marriage penalty relief provision contained in the Economic Growth and Tax Relief Reconciliation Act (EGTRRA) of 2001 to identify the ETI. In his study, the change in joint taxable income in response to the provision implies a joint elasticity ranging from 0.2 to 0.3.

Gruber and Saez (2002) is also one of the first studies to obtain elasticity estimates for both broad and taxable income.[12] At 0.12, their elasticity estimate for broad income is notably smaller than the corresponding estimate for taxable income, suggesting that much of the taxable income response comes through deductions, exemptions and exclusions. This is confirmed by the fact that the estimated ETI is much larger for itemisers (i.e. those taxpayers who make use of itemised deductions) than non-itemisers.

Giertz (2007) comes to a similar conclusion, but from a different angle. **The author investigates the change over time in estimated elasticities and relates this change to the availability of exemptions and deductions.** Applying the methods of Gruber and Saez (2002) to larger panel data sets between 1979 and 2001, Giertz (2007) obtains an estimated ETI for the 1990s of 0.2, or about half the size of the estimate for the 1980s. There is much less of a difference in the elasticity of broad income (0.15 in the 1990s as opposed to 0.12 for the 1980s). The estimates are consistent with the results from Kopczuk (2005) which suggests that the availability of deductions and exemptions matter in determining the ETI. Kopczuk (2005) notes that the fraction of taxpayers choosing to itemise was approximately 25% lower in the early 1990s compared to the mid-1980s, which could account for the lower ETI estimates for the later decade. This confirms the importance of the availability of deductions, exemptions and exclusions in the measurement of the income response to changing tax rates.

In general, estimated values of the ETI are lower in the 1990s (identified off of the 1990 and 1993 tax reforms) than in the 1980s (identified off of the 1981 and 1986 reforms). Apart from the availability of deductions and exemptions, there are two broad explanations for this pattern of results (Saez et al., 2012). The first is that there is no a priori reason to expect that the ETI is a universal parameter in the first place. Leading proponents of this view are Slemrod and Kopczuk (2002) who argue that the ETI is not a structural parameter, but is in fact a function of individual preferences as well as the breadth of the tax base and tax enforcement parameters (Kopczuk (2005) and Giertz (2007), among others, find evidence in support of this hypothesis). The second explanation (as proposed by Giertz, 2007) is that the model fails to adequately control for exogenous income trends. Rising income inequality could bias ETI estimates upward if tax rates fall and vice versa when rates increase. Additionally, the models fail to capture important types of income shifting, such as the shifting between the corporate and personal income tax base.

In fact, while the studies listed above attempt to control for mean reversion and secular income trends in a more nuanced way, they likely still underestimate the elasticity of taxable income. Weber (2014) shows that mean reversion prevents most estimators employed in the literature from obtaining consistent estimates of the ETI. Before estimating the ETI empirically, Weber (2014) examines a wide range of instruments, both theoretically and empirically, in order to determined which instruments will in fact be exogenous. The author uses the Michigan IRS Tax Panel data set for the years 1979 to 1990 to derive a consistent baseline ETI estimate of around 0.86, with the corresponding elasticity of broad income estimated at 0.475, both of which are somewhat higher than earlier estimates.

Most of the post-Feldstein literature focuses on legislated tax changes. However, several studies also investigated unlegislated changes and their impact on taxpayer behaviour. Saez (2003) examines responses to tax increases resulting from bracket creep between 1979 and 1981. During those years, inflation was around 10% a year but tax brackets were fixed in nominal terms. Bracket creep was a major issue in the late 1970s in the US, suggesting that taxpayers might have been aware of its effects. Saez regresses the one-year change in the log of income against the instrumented log change in the net-of-tax rate, along with base-year income controls and dummies for filing and itemization status, and finds a statistically insignificant overall ETI estimate of 0.31. For gross income, the estimated elasticity is 0.18 and insignificant, but when itemizers are considered separately, the estimate rises to 0.34 and is significant. This again suggests that itemizers are more responsive to tax changes.

[12] Broad income is defined as income before deductions.

Finally, Saez (2010) develops a method to estimate the compensated elasticity of reported income with respect to tax changes using bunching evidence around the kink points in the US federal income tax schedule. In the standard model, the amount of bunching should be proportional to the elasticity of the net-of-tax rate change at the kink (Saez et al., 2012). Saez (2010) finds evidence of bunching at the threshold of the first tax bracket where tax liability starts, especially in the 1960s. However, there is little evidence of bunching at other kink points in the tax schedule, even where the jumps in marginal tax rates are substantial.

2.4.2 ETI in other countries

While the bulk of research on the ETI has focused on the United States, some studies have investigated Canada and a select number of European countries. Importantly, given the fact the the ETI depends critically on the particular aspects of a country's tax system (such as the tax base definition, avoidance opportunities and the enforcement regime), there is no reason to expect the ETI to be the same across countries.

The 1988 Canadian tax reform raised statutory marginal rates for the two lowest income groups and lowered them for those falling in the top nine brackets, while broadening the tax base by either eliminating deductions and exemptions or converting them to non-refundable tax credits (Saez et al., 2009). Sillamaa and Veall (2001) use a panel of taxpayers who filed in both 1986 and 1989 and employ a method similar to Auten and Carroll (1999) to estimate the behavioural response of taxpayers to the reform. They estimate an ETI of 0.14 for those aged 25 to 61 and 0.27 for those older than 64. The corresponding estimates for broad income are 0.25 and 0.29 respectively. Similar to many of the U.S. studies, they report much higher elasticities for the upper income groups. Interestingly, their gross income elasticities are higher then their taxable income elasticities, in contrast the the usual finding based on U.S. data. This finding is possibly related to the fact that deductions in Canada are much more limited than in the United States.

Saez and Veall (2005) use aggregated-time-series income-share methods to analyse the effect of income taxation on the top income shares in Canada. They regress the log income share against the log net-of-tax rate for both the top 1% and top 0.1% of taxpayers, yielding elasticity estimates of 0.83 and 0.96 respectively. When they also include controls for the income share of the top 1% (or 0.1%), the elasticity estimates fall to 0.48 and 0.30.

In the United Kingdom, several studies have investigated the response to the Thatcher tax cuts. During the Thatcher administration, the top marginal tax rate on earnings fell from 83 % to 60% in 1979, and then dropped further to 40% in 1988. Dilnot and Kell (1988) first analysed the 1979 top rate cut. They find that the share of tax income from the top 1% of income earners stayed constant between 1978 and 1985 despite the tax rate cut, implying that top incomes must have increased following the rate cut. Brewer et al. (2010) perform a top income share time series analysis from 1962 to 2003. They find that the top 1% income share was falling in the United Kingdom until 1978, and started increasing precisely in 1979. Time series regressions of the log of the top 1% income share on the log net-of-tax rate produced significant elasticity estimates of around 0.5, even when controlling for changes in inequality.

Using a method similar to Gruber and Saez (2002), Kleven and Schultz (2014) use a panel of tax return data since 1980 to estimate the behavioural responses to the various income tax reforms implemented in Denmark between 1984 and 2003. Importantly, they have access to the full universe of Danish tax filers, generating a very large longitudinal dataset. They find that population-wide elasticities are modest compared to other studies (in the order of 0.2 to 0.3). However, similar to other studies, they show that elasticities are monotonically increasing in income level, with estimated elasticities 2 to 3 times larger in the top quintile of the income distribution than in the bottom.

Chetty et al. (2011) also use population tax return data from Denmark and estimate the ETI using bunching evidence around kink points, building on the method developed by Saez (2010). They developed a new method for estimating the long-run elasticity in an environment where adjustment costs attenuate the short-run response to tax reforms. Their results suggest that adjustment costs lead to heterogeneity in the size of the short-run elasticity response, depending on the size of the reform, which need to be taken into account when estimating the long-run elasticity. The long-run estimate is often of more interest for policy making than the short-run elasticities.

Selén (2002) and Hansson (2007) both estimate taxable income responses to Sweden's 1991 tax reform, employing somewhat different methodologies and relying on different data sets. Both studies employ difference-in-differences methodologies based on Feldstein (1995) and Moffitt and Wilhelm (2000), while Selén (2002) also employs a number of variations on Gruber and Saez (2002). Hansson (2007) uses a similar approach, but relies on just one pair of years. Importantly, using only two years makes it impossible to adequately control for mean reversion and/or changes in income inequality, casting doubt on the robustness of the results (Saez et al., 2009). Be that as it may, Hansson (2007) reports ETI estimates for Sweden that ranges from 0.37 to 0.43, similar to other international studies. Selén (2002) reports estimated ETIs that are generally between 0.2 and 0.4, but reports an ETI of 0.53 when focusing on those taxpayers in the top two income brackets.

Gottfried and Schellhorn (2004) follow the approach in Gruber and Saez (2002) to estimate the taxable income responses to the German tax reforms of 1990, but use only one pair of years, 1988 and 1990. They report a compensated elasticity of 0.58 (and an income elasticity of zero). As mentioned above, using only two years might compromise identification in the presence of mean reversion and/or changes in income inequality.

Atkinson and Leigh (2008) analyse the top 1% income shares in New Zealand from 1921 to 2005. They regress the top 1% income share on one year lags of the net-of-tax rate and several controls, finding an overall ETI estimate of 0.41.

Other studies focus on France (Piketty, 1999), Finland (Pirttilä and Selin, 2006), Poland (Kopczuk et al., 2012) and Russia (Gorodnichenko et al., 2009), to name a few. In general, ETI estimates are in the 0.2 to 0.8 range.

3 Estimating the ETI for South Africa

3.1 South African Personal Income Tax System and Tax Reforms

Income tax in South Africa was first introduced in 1914 with the introduction of the Income Tax Act of 1914 (Act no. 28 of 1914). The act has gone through numerous amendments, with the act presently in force being the Income Tax Act of 1962 (Act no. 58 of 1962). Several other laws govern various aspects of the tax code. These include, among others, the Value-Added Tax Act (1991) and the Employment Tax Incentive Act (2013).

Normal income taxes are levied on all individuals in the form of an annual tax that is calculated by applying predetermined rates to an individual's taxable income. Taxable income is defined as gross income (salary and wage income, business income, interest income, and sundry income such as dividend income and annual payments, including bonuses) less exempt income and allowable deductions.[13]

Since 1994, the personal income tax system has undergone numerous changes. Against the backdrop of a rising personal income tax burden in the 1980s, the Katz Commission was tasked with a comprehensive review of the South African tax system. The commission's first interim report was released in 1994 (Katz Commission, 1994). This was followed by another nine reports between 1994 and 1999.

[13] In the case of South Africa, exemptions include, among others, interest income earned from a South African source up to a certain threshold. Itemised deductions include, among others, current and arrear pension fund contributions, current and arrear retirement annuity fund contributions, medical and disability expenses, and donations.

Based on the recommendations of the Katz Commission (and other role-players, including National Treasury), the South African income tax system was changed in many respects over the last 20 years. The number of personal income tax brackets was reduced from 10 to 6; the child rebate was scrapped; the individual is now the unit of taxation; the primary rebate was increased annually to compensate for inflation and to retain progressivity; capital gains tax was introduced and the secondary tax on companies was replaced by a dividends tax.

A brief overview[14] of the historic development of income tax reform in South Africa since the transformation shows that in 1993/1994, the number of income bands was limited to 9 and individuals were taxed at a minimum and a maximum marginal rate of 17% and 43%, respectively. At the time they were also taxed differently according to marital- and gender status. In 1995/1996 only the highest marginal rate increased to 45% with the number of income bands stretched to 10 while gender differentiation was discontinued. In the 1997/1998 fiscal year the lowest marginal rate increased to 19% and the income bands decreased to 7. Additionally, the tax free threshold was adjusted in order to differentiate between individuals younger than 65 years and older than 65 years, while the child rebate was removed. In 1998/1999 the income bands were reduced to 6 and in 2000/2001 the top marginal rate decreased to 42%, while the lowest marginal rate was reduced to 18%. Additionally, South Africa moved from a source-based system of taxation to a residence-based one with effect from 1 March 2001.

Capital gains tax was introduced in October 2001 and is currently still in effect. Capital gains tax is effectively charged by adding a percentage of the increase in value of an asset that was disposed of for more than its base cost to the taxpayer's taxable income and taxing income at the relevant marginal rate. For individuals, deceased estates and special trusts, 40% of the net gain is added to their taxable income. For companies, close corporations and trusts, 80% is added.[15]

The top marginal tax rate was reduced to 40% in 2002/2003 and remained there until 2015/2016 when it was increased to 41%. **In 2017/18 a new top income tax bracket with a marginal rate of 45% was introduced.** In 2011/2012 a further rebate was introduced for individuals older than 75. Finally, a dividends tax in the form of a 15% withholding tax on dividend income received from companies and closed corporations was introduced in April 2012. The dividends tax replaced the so-called secondary tax on companies (STC).

In July 2013, the Davis Tax Committee, chaired by Judge Dennis Davis, was appointed to review the South African tax policy framework. The main goal of the committee is to assess the tax policy framework and its role in supporting the South African government's objectives of inclusive growth, employment, development and fiscal sustainability. Several interim reports have been published since 2013 with recommendations relating to value added tax (VAT), estate duties and mining and carbon taxes.

Several implications of the (changing) tax code described above are worth noting in the context of the present study. Firstly, the sample used in this study does not include any major PIT tax reforms. Between the 2008/9 and 2012/13 tax years, marginal tax rates remained unchanged, while the reforms that were implemented (including the switch from the STC to the dividends tax and the introduction of the tertiary rebate) were insignificant in the sense that they are unlikely to have induced a major behavioural response on the part of individual taxpayers.

The lack of major tax reforms complicates the identification of the ETI. During the sample period in question, there was no large, policy-induced variation in marginal tax rates of the kind that is often used to identify and estimate the ETI. As such, we will follow Saez (2003) in using the phenomenon of 'bracket creep' to identify the ETI.

Over our sample period, National Treasury adjusted individual tax brackets in order

[14]See van Heerden (2013).

[15]For the 2017/2018 tax year, this implies that capital gains are effectively taxed at a maximum rate of 18% for individuals (based on a top marginal tax rate of 45%).

to compensate for the effects of inflation.[16] The extent to which official bracket adjustments compensate for inflation depends critically on the measure of inflation that is used. Popular inflation measures include the consumer price index (CPI) and the GDP deflator. However, these measures might underestimate the extent to which nominal incomes change from year to year. If nominal wages increase in excess of official inflation measures, taxpayers near the top end of an income tax bracket might still be pushed into the next bracket despite the announced bracket adjustments. This is particularly true in the case of South Africa where, in the past, wage adjustments have often been in excess of consumer price inflation.

Table 1 shows the average effective bracket adjustment for each tax year in question along with two inflation measures (based on the CPI and the GDP deflator) and a measure of economy-wide wage inflation. Apart from 2008/09, bracket adjustments were generally in line with inflation as measured by the CPI and GDP deflator. However, as mentioned above, when studying the phenomenon of 'bracket-creep' the relevant inflation measure is nominal wage inflation. When comparing average bracket adjustments to economy-wide wage inflation, we see that, apart from 2008/09 and 2012/13, bracket adjustments were not in line with nominal wage inflation. This suggests that a significant proportion of taxpayers likely migrated to higher tax brackets and, as a consequence, faced a higher marginal tax rate in the following tax year. It is this phenomenon of 'bracket creep' that will be used to identify the ETI.

Table 1: Bracket adjustment and inflation

Tax year	Average bracket adjustment	CPI inflation	GDP deflator	Economy-wide wage rate
2008/09	8.4%	9.9%	8.9%	7.7%
2009/10	7.7%	7.1%	7.5%	11.6%
2010/11	5.3%	4.3%	6.3%	14.8%
2011/12	6.1%	5.0%	6.5%	7.9%
2012/13	6.5%	5.3%	5.3%	6.2%

Source: National Treasury, Statistics South Africa, South African Reserve Bank. Inflation rates calculated as year-on-year % changes in CPI, the GDP deflator, and the total economy-wide wage rate.

As mentioned in Section 2.3.1, taxpayers might not be fully aware of small changes to the tax code and, as such, any estimated behavioural response in reaction to such changes will be muted. While income tax brackets were not fully adjusted for wage inflation, they did not remain constant in nominal terms as was the case in the US in the period under consideration in Saez (2003). As a result, there was relatively little variation in marginal tax rates between 2009 and 2013. This implies that the estimated ETI in Section 3.4 likely underestimates the true behavioural response to a change in marginal tax rates in the case of South Africa.

Finally, it is worth noting that the tax code described above (i.e. graduated income tax schedule, allowable deductions, low-income tax rebates, exemptions etc.) ultimately results in a highly progressive income tax system, consistent with the idea that the wealthy should contribute a greater proportion towards supporting the State than the poor. According to the 2015 Tax Statistics (published jointly by SARS and the National Treasury), 58% of total assessed personal income taxes came from the 9.7% of individual taxpayers who have taxable incomes of R500 000 or more. The progressivity of the system is also clear when looking at corporate income taxes, where almost 64% of tax comes from as little as 0.9% of the large companies that are profitable enough to pay tax at all.

[16]National Treasury uses historical CPI inflation, i.e. previous calendar year CPI inflation, as a base for the bracket adjustment calculations, tailoring the adjustment according to budgetary requirements. For example, CPI inflation measured 6.3% in 2016, but tax brackets were only adjusted by 1% in 2017/18 due to additional revenue requirements.

The progressive nature of the tax code implies that most of the action in terms of the behavioural response of individual taxpayers to changing marginal tax rates will be concentrated at the top of the income distribution. A number of empirical studies have confirmed this fact internationally, with less evidence of any response for the low, middle and upper-middle income classes (Saez et al., 2012). This suggests that it might be useful to disaggregate the estimated aggregate ETI by income group in order to identify possible differentiated behavioural responses. Such differentiated response have implications for the formulation of optimal tax policy.

3.2 'Bracket Creep', Data and Descriptive Statistics

3.2.1 'Bracket Creep'

As mentioned above, the sample period in question did not contain any large tax reforms of the kind that is often used to identify and estimate the ETI. We will follow Saez (2003) in using the phenomenon of 'bracket creep' to identify the ETI.

Saez (2003) examines responses to tax increases resulting from 'bracket creep' between 1979 and 1981 for the US. During those years, inflation was around 10% a year but tax brackets were fixed in nominal terms, resulting in taxpayers graduating to higher income brackets despite real income remaining constant (or even declining) between years.

In contrast, in South Africa nominal tax brackets have regularly been adjusted. However, these adjustments have not always fully accounted for the effects of inflation. By not fully adjusting for the effects of inflation, tax revenues can increase automatically as taxpayers graduate to higher income tax brackets.

As mentioned above, the extent to which bracket adjustments compensate for inflation depends critically on the measure of inflation that is used. Popular inflation measures include the consumer price index (CPI) and the GDP deflator. However, these measures might underestimate the extent to which nominal incomes change from year to year. As such, we follow Gruber and Saez (2002) in calculating an inflation measure based on the average growth of gross or broad income for each of the years in the sample.[17] This measure more accurately reflects the actual change in nominal incomes from year to year. Should official bracket adjustments not fully compensate for (above inflation) nominal income growth, the 'bracket creep' effect will be more pronounced.

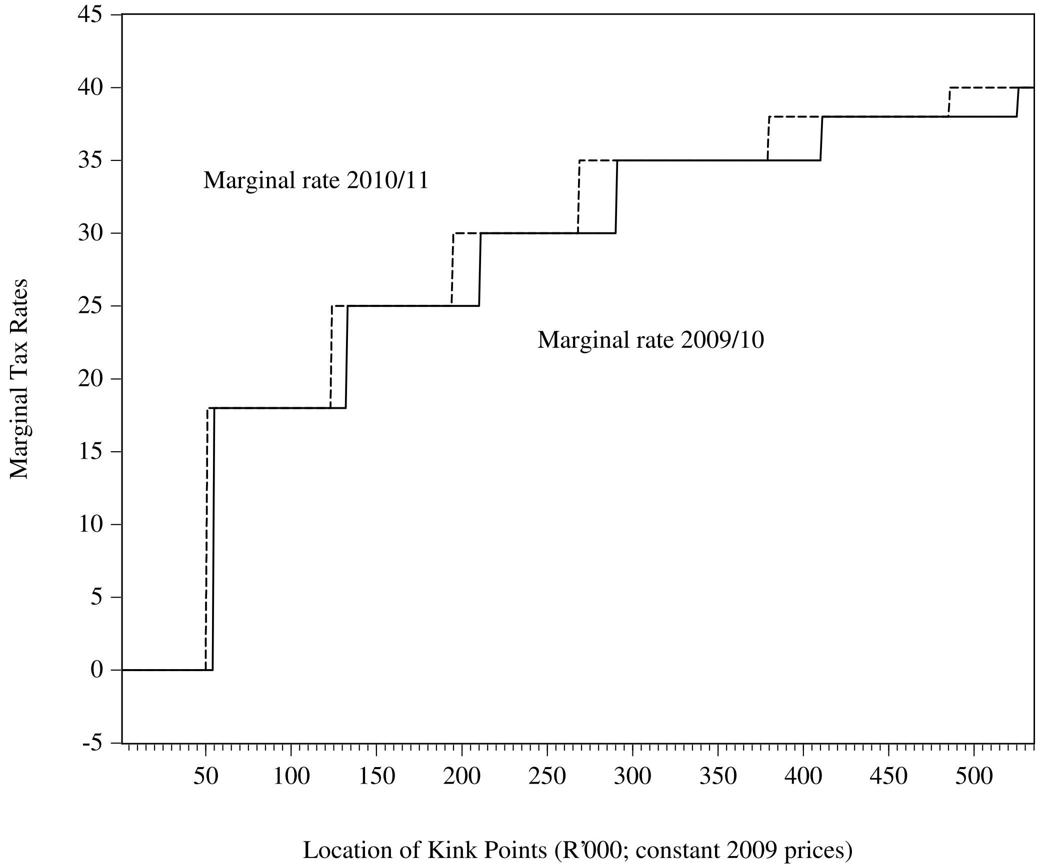
Figure 1 displays the effect of our chosen inflation measure on marginal tax rates between 2009/10 and 2010/11.[18] Marginal rates as a function of before-tax *real* income is plotted for the 2009/10 (solid line) and 2010/11 (dotted line) tax years. If taxable income stays constant in *real* terms, then some taxpayers will face a higher marginal rate in year two. Apart from the lowest income bracket, the changes in marginal rates are not very large (between 2% and 5%). This is a far cry from the large reforms often studied in this literature.[19] However, the 'bracket creep' phenomenon does provide for some variation in tax rates, even in an environment where there was no legislated changes in marginal tax rates.

[17] See Section 3.2.2 for more detail.

[18] Our measure of inflation was largest across this pair of years, resulting in a clear illustration of the 'bracket creep' effect.

[19] For example, the 1986 Tax Reform Act in the US led to a decrease from 50% to 28% in marginal rates for very high income earners.

Figure 1: Shift in marginal tax rates between 2009/10 and 2010/11



3.2.2 Data and Descriptive Statistics

The analysis presented here uses confidential tax return data, made available for research purposes by SARS and the National Treasury. The available sample runs from 2008 to 2012.[20] The full dataset contains most items found on individual Employee Tax Certificate (IRP5/IT3(a)) and/or Income Tax Return (ITR12) forms. This includes information on various aspects of personal income taxation, including information on labour, business and investment income, deductions, exemptions and allowances, taxable income, and total tax liability. The full dataset also contains information on assessed taxes, i.e. income tax returns that were processed and assessed by SARS.

It was decided to focus the analysis on those taxpayers who submitted tax returns between 2008 and 2012 that were, in turn, assessed by SARS. Importantly, focusing on assessed taxes alone significantly reduces the number of observations available for the analysis. Between 2007 and 2011, individual taxpayers in South Africa were not liable to file a tax return if their employment income was less than R120 000 per annum, provided that the income was derived from employment only (i.e. labour income as recorded on the IRP5 form). This limit was increased to R250 000 in 2012.

International evidence points to the fact that most of the action in terms of behavioural responses to changing marginal tax rates is concentrated in higher income groups and among those taxpayers with itemised deductions. Given that the assessed tax database includes only those individuals with relatively high personal income levels and/or itemised deductions, focusing on this dataset will still allow us to obtain an estimate of the ETI that reflects the broad behavioural response of the South

[20] In truth, the sample runs from the 2008/09 tax year to the 2012/13 tax year. In South Africa, the tax year runs from 1 March to 28 February of the following year. For notational convenience we will refer only to the *calendar* year which comprises the largest proportion of the relevant tax year, i.e. 2008/09 becomes 2008, 2009/10 becomes 2009, etc.

African taxpayer to changing marginal tax rates. Additionally, given the relatively small changes in marginal rates induced by the 'bracket creep' phenomenon, it is unclear whether those individuals with lower (single-source) income and no itemised deductions would have responded by altering their labour supply decisions.

Finally, according to official statistics published by SARS, assessed taxes on average made up over 85% of total personal income taxes between 2008 and 2012. This suggests that any behavioural change for this group of taxpayers will have significant revenue implications.

To facilitate estimation of the ETI, a balanced panel was constructed using the assessed dataset by selecting only those individuals present in each of the years in the sample, resulting in just over 3 million observations per year.[21] Table 2 presents means and standard deviations of the relevant income data (in real terms) for the constructed panel over the entire sample period (2008 - 2012).

Table 2: Summary statistics: Balanced panel
(2008-2012; constant 2008 prices)

	Mean	S.D.
Broad income	R225 557	R366 784
Normal income	R190 364	R238 024
Other income#	R77 498	R447 586
Investment income#	R54 017	R309 056
Business income#	R87 914	R333 824
Capital gains income#	R102 304	R818 943
Taxable income	R202 274	R351 935
Itemisers	83%	
Number of observations	15 041 970	

#Statistics for other income, investment income, business income and capital gains income calculated by only taking into account those individuals with recorded values for the respective income sources.

We use two types of definitions for income: gross or broad income, and taxable income. Broad income is an extensive definition of income that is consistent across all the years in the sample. It includes most of the items that are summed to arrive at total income (as recorded on IRP5/IT3(a) and ITR12 forms), namely wage income, investment income, business income, fringe benefits, allowances, etc.[22] We follow the international literature in excluding capital gains income from the broad income calculation due to the special tax treatment afforded to capital gains. Taxable income is defined as broad income less exemptions and other allowable deductions.

According to Table 2, average broad income in the sample equals about R225 000 and average taxable income equals around R202 000, measured in constant 2008 values. Eighty-three percent of the individuals in the dataset recorded itemised deductions.

Following Gruber and Saez (2002), the empirical strategy is to relate changes in income between pairs of years to the change in marginal tax rates between the same pairs of years. The time length between year 1 (the base year) and year 2 can be of 1, 2 or 3 years. In our basic specification, we follow Gruber and Saez (2002) and Feldstein (1995) in setting the time length equal to 3 years. Therefore, we relate year 2011 to year 2008, and year 2012 to year 2009. These two differences are stacked to obtain a single dataset of about 6 million observations. Given the fact that the identification strategy employed here relies on what is essentially unobserved tax rate changes, the time lag allows for a more complete behavioural response as taxpayers become aware of the tax implications of 'bracket creep'.

[21]**Before creating the panel, the dataset was cleaned by removing one outlier, as well as selecting only on those individuals with recorded values for gross and taxable income. No transformations were applied.**

[22]A detailed description of the construction of the income measures is provided in the Appendix.

As mentioned above, our inflation measure is based on the average growth of broad income for each of the years in the sample. Taking 2008 as the base year index (equal to 100), the incomes for years 2009 to 2012 have been deflated using the following indices: 109.61, 124.59, 137.70 and 150.52.

3.3 Empirical Strategy

We use the basic model discussed in Section 2.1 to derive a regression specification. Recall that the basic model boils down to estimating e^c in Equation (3), repeated here for convenience.

$$\frac{dz}{z} = -e^c \frac{d\tau}{1-\tau} + \eta \frac{dR - zd\tau}{z(1-\tau)}$$

where z is real income, e^c is the compensated elasticity of taxable income, τ is the relevant marginal tax rate, η is the income effect and R is virtual income. This equation displays the behavioural response in reported income induced by a small tax change. Following Gruber and Saez (2002), this equation can be estimated by replacing z by z_1 (year 1 income), $\mathrm{d}z$ by $z_2 - z_1$ (change in income between year 1 and year 2), $\mathrm{d}\tau$ by $T_2'(z_2) - T_1'(z_1)$ (the change in marginal tax rates), and $\mathrm{d}R - z$ by $[z_2 - T_2(z_2)] - [z_1 - T_1(z_1)]$ (the change in after-tax income).[23]

Using the log-log specification and replacing $\mathrm{d}z/z$ by $\log(z_2/z_1)$, $-\mathrm{d}\tau/(1-\tau)$ by $\log[(1-T_2')/(1-T_1')]$, and $(\mathrm{d}R - z\mathrm{d}\tau)(z(1-\tau))$ by $\log[(z_2 - T_2(z_2))/(z_1 - T_1(z_1))]$, we obtain the following regression specification:

$$\log(z_2/z_1) = e\log[(1-T_2')/(1-T_1')] + \eta\log[(z_2 - T_2(z_2))/(z_1 - T_1(z_1))] + \epsilon \tag{11}$$

where e is the compensated elasticity parameter, η is the income effect parameter, z_i is *real* income in year i, T_i' is the marginal tax rate in year i and $T_i(z_i)$ is the tax liability in year i.

Let's ignore the income effect for now (i.e. set $\eta = 0$). The term capturing the tax rate change $\log[(1-T_2')/(1-T_1')]$ is correlated with ϵ because if there there is a positive shock to income (i.e. $\epsilon > 0$), then the tax rate increases mechanically due to the progressive nature of the income tax system. Therefore, an OLS estimate of (11) will produce a biased estimate of the behavioural elasticity.

We follow Gruber and Saez (2002) and Saez (2003), among others, in building an instrument for this variable by computing T_p', the marginal tax rate that an individual would face in year 2 if their *real* income did not change from year 1 to year 2. That is, $T_p' = T'(z_p)$ where z_p is predicted taxable income in year 2, which is taxable income in year 1 expressed in year 2 rands. The natural instrument for $\log[(1-T_2')/(1-T_1')]$ is therefore $\log[(1-T_p')/(1-T_1')]$, or the predicted change in the log net-of-tax rate if real income does not change between year 1 and year 2.

Columns (1) and (4) in Table 3 provides information on the value of the instrument for each pair of years in our sample for the six different tax brackets (defined based on base year taxable income).[24] By construction, the instrument is negative for a tax rate increase and positive for a tax rate cut. Over both pairs of years and across all income groups the values are negative, reflecting the fact that 'bracket creep' pushed most taxpayers into higher tax brackets. The value for the instrument is zero in the top tax bracket, suggesting that those who were in the top bracket in year 1 remained there in year 2. The fact that there are no positive values for the instrument suggests that nominal tax brackets were not fully adjusted for (our measure of) inflation over the sample period.

[23]Here, $z(1-\tau)$ is approximated by $z - T(z)$

[24]Baseline results are presented for the 3-year difference case. The implications of different lengths of differences is discussed in Section 3.5.

Table 3: Variation in after-tax shares and mean reversion in income

		2008-2011			2009-2012		
		Instrument	Difference in log of taxable income	Difference in log of broad income	Instrument	Difference in log of taxable income	Difference in log of broad income
		(1)	(2)	(3)	(4)	(5)	(6)
Tax Bracket[#]	1	-0.01	0.17	0.19	-0.02	0.17	0.12
		[0.03]	[0.71]	[0.66]	[0.03]	[0.69]	[0 .62]
		1 218 858	1 218 858	1 218 858	1 156 133	1 156 133	1 156 133
	2	-0.02	-0.04	0.01	-0.02	0.00	-0.01
		[0.03]	[0.40]	[0.36]	[0.03]	[0.41]	0.36
		861 336	861 336	861 336	877 650	877 650	877 650
	3	-0.03	-0.10	-0.05	-0.03	-0.05	-0.06
		[0.04]	[0.45]	[0.40]	[0.04]	[0.45]	[0.40]
		367 613	367 613	367 613	394 910	394 910	394 910
	4	-0.02	-0.15	-0.11	-0.02	-0.09	-0.09
		[0.02]	[0.49]	[0.43]	[0.02]	[0.49]	[0.43]
		254 070	254 070	254 070	270 487	270 487	270 487
	5	-0.02	-0.18	-0.15	-0.02	-0.13	-0.13
		[0.02]	[0.52]	[0.46]	[0.02]	[0.51]	[0.46]
		121 788	121 788	121 788	120 994	120 994	120 994
	6	0.00	-0.29	-0.26	0.00	-0.21	-0.21
		-	[0.65]	[0.59]	-	[0.63]	[0.58]
		184 729	184 729	184 729	188 220	188 220	188 220

[#]Taxpayers grouped into tax brackets based on base year taxable income (1 = bottom bracket, 6 = top bracket). Means, standard deviations (in square brackets) and number of observations are reported. Taxable income and broad income expressed in constant 2008 prices.

Finally, note that while there is little variation in the mean value of the instrument between income groups (or tax brackets), there is substantial heterogeneity within groups, as illustrated by the relatively large standard deviations.

According to Gruber and Saez (2002), running an instrumental variables (IV) regression of (11) might also lead to a biased estimate of the elasticity if ϵ is correlated with z_1. This relates to the earlier discussion on mean reversion and a changing income distribution. High incomes in year 1 tend to be lower in following years (mean reversion), producing a negative correlation between ϵ and z_1. On the other hand, should the income distribution widen, there will be a positive correlation between ϵ and z_1. If ϵ is correlated with z_1, then the instrument (which is also a function of z_1) will also be correlated with the error term, producing biased estimates. It is for this reason that many studies in the literature follow Auten and Carroll (1999) in including lagged income as a control, either in a simple way by including $\log z_1$ or in a richer way by including polynomials or splines in year 1 incomes.

There is definite evidence of mean reversion in our sample, both in terms of taxable income and broad income. Columns (2)-(3) and (5)-(6) in Table 3 shows the mean values of our dependent variables for each pair of years in our sample for the different income groups. There is mean reversion at both ends of the income distribution. Changes in incomes are high and positive for low incomes, whereas the change in incomes become in general negative for high income earners. This complicates the estimation of the elasticities, particularly at the low and high ends of the income distribution. For this reason, we include several base year income controls in our final regression specification.

Turning to the income effect, the term $\log[(z_2 - T_2(z_2))/(z_1 - T_1(z_1))]$ will also be correlated with ϵ for the same reasons discussed above. However, in much of the literature the income effect is assumed away (i.e. $\eta = 0$) since those studies that do estimate income effects often find them to be small and insignificant.

In the current study we have additional motivation for omitting income effects from our base specification. In the presence of 'bracket creep', the income effect term affects

those individuals who experience a change in marginal rates (treatments) and those individuals who do not experience a change in marginal rates (controls) in approximately the same way (Saez, 2003). Therefore, this additional income effect can be incorporated into the error term.

To see why, recall that $dR - zd\tau$ in equation 3 is the change in after-tax income due to the tax change for a given before-tax income z. Because of 'bracket creep', this quantity is piece-wise linearly but continuously increasing in income (i.e. flat over control regions and linearly increasing over treatment regions) and thus affects treatments and controls in approximately the same way. Intuitively, at any given kink in the tax schedule the increase in tax liability due to 'bracket creep' is approximately the same for both treatments and controls, but the change in marginal tax rates is different. As a result, the difference in the behavioural response between the two groups is almost exclusively due to pure substitution effects. This implies that, in this framework, the estimated elasticity e when $\eta = 0$ is, in fact, equal to the compensated elasticity, e^c, our parameter of interest. Therefore, this additional income effect term can be incorporated in the error term. The dependence of this error term on income is controlled for by the functions in $taxinc_1$.

The precise regression framework is the following:

$$\begin{aligned}\log(z_2/z_1) = \alpha_0 + e^c \log[(1-T'_2)/(1-T'_1)] + \alpha_1 \log z_1 + \alpha_2 f(taxinc_1) \\ + \sum_{i=1}^{10} \alpha_{3i} SPLINE_i(z_1) + \sum_j \alpha_{4j} YEAR_j + \beta item + \epsilon\end{aligned} \tag{12}$$

where z_i is *real* income in year i (either taxable income or broad income), T'_i is the marginal rate in year i and e^c is the parameter of interest, i.e. the compensated elasticity. $YEAR_j$ denote base year dummies and *item* is a dummy variable for being an itemiser in year 1 (i.e. individuals with recorded itemised deductions). $SPLINE_i$ is a ten-piece spline in base year income (included as control in some specifications), while $f(taxinc_1)$ are smooth functions in (nominal) base year income (polynomial terms in $taxinc_1$). Polynomial terms are added until the elasticity estimate is stabilised (two terms are sufficient in most cases).

The dataset is constructed by stacking differences in income across individuals and years. Equation (12) is then estimated by simple 2SLS, the first stage being:

$$\begin{aligned}\log[(1-T'_2)/(1-T'_1)] = \log[(1-T'_p)/(1-T'_1)] + \theta_1 \log z_1 + \theta_2 f(taxinc_1) \\ + \sum_{i=1}^{10} \theta_{3i} SPLINE_i(z_1) + \sum_j \theta_{4j} YEAR_j + \gamma item + \epsilon\end{aligned} \tag{13}$$

where $\log[(1-T'_p)/(1-T'_1)]$ is used as an instrument for $\log[(1-T'_2)/(1-T'_1)]$. The first stage of the regression is very strong, with the F-statistics for the coefficient of the tax rate instrument always in excess of the relevant critical value.[25]

Finally, by construction of the dataset and due to the fact that we stack observations from two pairs of years to form our estimates, we are using multiple observations on the same individuals. If there is individual-specific correlation in how income changes over time, standard 2SLS will understate the associated standard errors. Because of this, all reported standard errors are corrected for intra-personal correlation.

[25]First stage results reported in the Appendix.

3.4 Basic Results

The basic results from the 2SLS regressions are reported in Table 4 (standard errors in parentheses). The table has 8 columns, expressing three alternative methods for dealing with mean reversion and/or income distribution changes for the two income concepts. In columns (1) and (5) we do not include any controls. In columns (2) and (6) we include log income, as in Auten and Carroll (1999), among others. Coulmns (3) and (7) include polynomial terms in base year income as additional controls, as in Saez (2003). Finally, we follow Gruber and Saez (2002) by including a 10-piece spline in base year income to allow for non-linearities in the widening of the income distribution. **There is no reason to expect that the two main sources of bias in estimating the ETI (mean reversion and income distribution changes) operate linearly, particularly in combination with each other. As such, one specification includes richer base period income controls (splines) in attempt to address the possible non-linearity.**

Results are presented for both definitions of income, broad and taxable. All regressions are weighted by income to reflect the relative contribution to total revenue, while at the same time facilitating the analysis of optimal taxation.[26]

The results reflect substantial sensitivity to controlling for income. For the models in Table 4 that exclude any form of control for mean reversion and/or income distribution changes we obtain large and, in the case of taxable income, wrong-signed elasticity estimates. Once log income is included, the results change dramatically. For broad income, the elasticity falls to 0.17, while for taxable income the elasticity becomes a positive 0.30. This estimate lies close to the mid-point of the post-Feldstein literature discussed above. Log income itself has a highly significant negative coefficient, suggesting that on average mean reversion dominates income dispersion in our sample period. Coefficients on the polynomial terms in base year income, while small, are highly significant. Including the polynomial terms results in slightly smaller elasticity estimates.

These specifications assume that any change in the income distribution is a simple linear function of lagged income. We can weaken this assumption by including a 10-piece spline in lagged income. **The estimated coefficients are all significant, suggesting that there is some evidence of non-linearity in the change in income distribution.** Including the spline results in only a small change to the elasticity in taxable income (up to 0.31), but a more significant change in the case of broad income (up to 0.24). For both broad and taxable income, the splines are highly negative at the bottom of the income distribution, perhaps reflecting worsening income prospects for low income groups over this period. The coefficients display a high degree of non-linearity throughout the rest of the income distribution.

It is clear from the results that there is substantial (and statistically significant) differences between the elasticity estimates for broad and taxable income. There are two sources of difference here. The first is purely mechanical - broad income has a larger base so that any rand response will result in a smaller elasticity. The second, which has been referred to above and within the ETI literature, is behavioural. Taxable income includes itemised deductions, which might respond to changes in taxes. This effect is clearly demonstrated when looking at the estimated coefficient on the itemiser dummies in Table 4. The coefficients are statistically significant and positive across all specifications, suggesting that a large share of the behavioural response to changing marginal tax rates takes place through itemised deductions.

Finally, the critical values reported final two rows of Table 3.4 suggest that we can reject the null of exogeneity with respect to the instrumented variable, namely $\log[(1-T_2')/(1-T_1')]$**. This, combined with the first stage results reported in the Appendix, suggests that the instrumental variables approach is appropriate.**

[26]As discussed in Gruber and Saez (2002) and Saez et al. (2012), the important parameters for the analysis of optimal taxation and deadweight burden are the elasticities weighted by income because the income response to a change in marginal rates is proportional to the elasticity times the income level.

Table 4: Basic elasticity results

	Taxable Income				Broad Income			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
e^c	-1.820***	0.304***	0.262***	0.314***	1.764***	0.167***	0.139***	0.235***
	[0.053]	[0.038]	[0.019]	[0.059]	[0.048]	[0.034]	[0.018]	[0.055]
$item$	0.008**	0.105***	0.097***	0.101***	-0.017***	0.096***	0.084***	0.092***
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.003]	[0.003]
$\log(z_1)$		-0.190***	-0.161***			-0.184***	-0.155***	
		[0.006]	[0.006]			[0.005]	[0.005]	
$taxinc_1$			-1.62e-08**				-1.66e-0.8***	
			[5.17e-0.9]				[4.74e-0.9]	
$taxinc_1^2$			2.22e-17**				2.26e-17**	
			[8.26e-18]				[7.54e-18]	
Spline 1^{st} decile control				-0.704***				-0.684***
				[0.003]				[0.003]
Spline 2^{nd} decile control				-0.133***				-0.059***
				[0.004]				[0.004]
Spline 3^{rd} decile control				-0.105***				-0.127***
				[0.012]				[0.010]
Spline 4^{th} decile control				-0.183***				-0.161***
				[0.013]				[0.013]
Spline 5^{th} decile control				-0.189***				-0.122***
				[0.015]				[0.009]
Spline 6^{th} decile control				-0.069***				-0.110***
				[0.020]				[0.013]
Spline 7^{th} decile control				-0.173***				-0.159***
				[0.009]				[0.006]
Spline 8^{th} decile control				-0.197***				-0.224***
				[0.013]				[0.013]
Spline 9^{th} decile control				-0.013				-0.003
				[0.022]				[0.022]
Spline 10^{th} decile control				-0.237***				-0.236***
				[0.013]				[0.013]
Observations	6 016 788	6 016 788	6 016 788	6 016 788	6 016 788	6 016 788	6 016 788	6 016 788
Adjusted R^2	0.255	0.065	0.078	0.069	0.250	0.093	0.105	0.089
Wu-Hausman test statistic	248 526***	371 822***	400 518***	363 206***	173 501***	282 371***	304 399***	289 339***
Robust F-statistic	8519***	40 039***	40 419***	16 244***	6767***	37 509***	32 233***	11 944***

Estimates of 2SLS regressions. *, **, *** reflect significance at the 5%, 1% and 0.1% levels respectively. Regressions weighted by income. All regressions include dummy variables for each base year.

The basic results recorded in Table 4 suggest that there is a sizeable response of taxable income to tax changes, **with a plausible range (given the controls) for the estimated elasticity of 0.26 to 0.31**. This is well within the range of the post-Feldstein literature.[27] On the other hand, the elasticity of broad income is much smaller. The gap can partly be explained by changes in itemisation and/or exemption behaviour.

It must be noted that the estimated elasticities reported in Table 4 likely underestimates the full behavioural response to a change in the marginal tax rate. While the 'bracket creep' phenomenon studied here provides for some variation in marginal rates over the sample period and allows for the use of a difference-in-differences estimator, the implicit tax changes were relatively small and

[27] As an additional sensitivity analysis, we follow Gruber and Saez (2002) and censor the change in log income at 5 so that the approximately 3000 observations who report changes in income ratios across the 2 years of more than 150 or less than 1/150 are censored. This avoids the influence of job shifting (to some extent) and/or large once-off increases/decreases in income. The results are not very sensitive to this restriction (see the Appendix for more details) and as such we proceed with the uncensored dataset.

unlegislated. The absence of any legislated and well-publicised tax reform over the sample period implies that the behavioural response to changing tax rates might not be fully captured in the current framework.[28]

Finally, our data precludes the identification of the source of the behavioural response. As mentioned above, the ETI aims to capture all possible behavioural responses to changes in income taxation, including labour supply choices and tax evasion/avoidance, in a single measure. The source of the behavioural response is of obvious importance for the formulation of policy. Unfortunately, evidence on labour supply elasticities with respect to changing tax rates, and/or tax evasion and avoidance, for the South African economy is scant.

There is some limited evidence on the responsiveness of labour supply to real wages. However, in general, there is little focus on the influence of wage growth (in particular) on labour supply in the South African context, with most authors focusing on the demographic drivers of labour force participation instead. What evidence there is is mixed. Von Fintel (2016) uses a district pseudo-panel, constructed from household surveys, to estimate the elasticity of labour demand, labour supply and unemployment with respect to wages and finds that labour supply is generally inelastic w.r.t. wages. Fedderke (2012) surveys the evidence on labour market rigidities and find that some studies imply (relatively) high labour supply elasticities, while others point to lower values. One problem is that there is little consensus on the correct way to estimate these elasticities. Be that as it may, the literature does suggest that elasticities are generally non-zero, suggesting some supply response to changing wages. If one considers the fact that migrating to a higher tax bracket can result in lower take-home pay (i.e. lower real income), at least some of the behavioural response to changing tax rates can be assumed to originate from a labour supply response.

Turning to tax evasion/avoidance, the evidence is spread thin. Most studies comprise surveys among different population groups w.r.t. their attitude toward tax avoidance strategies. For example, Oberholzer and Stack (2014) find that different population groups have different opinions on tax evasion/avoidance, but that, in general, 25% of respondents express the desire to keep all earned income (i.e. not pay taxes), while 13% believe many other people avoid paying taxes and so see nothing wrong with doing the same. Robinson and Gcabo (2007) find that 50% of respondents in their survey would consider cheating on their taxes under certain circumstances. Ross and McGee (2012) find that attitudes toward tax evasion differ by demographic variable, with higher income individuals being more amenable to tax avoidance strategies. The results on the acceptability of tax avoidance suggest that at least some of the behavioural response captured in the estimated ETI can be attributed to tax evasion/avoidance. However, the dataset precludes the estimation of the magnitude of the response.

3.5 Variations in Timing

We have followed previous literature in using a 3-year difference in computing both the change in taxable income and the change in after-tax shares. Using the framework above we can test the sensitivity of our results to a change in the difference window.

The exact implications of changing the window of observation is not clear. If individuals react slowly to tax changes, a longer window might increase the estimated elasticity. If, however, responses to

[28] **On the other hand, the particular economic circumstances of the period in question might have *increased* taxpayer sensitivity to changing tax rates. The sharp decline in economic activity in 2009, and subsequent recovery, likely induced behavioural responses on the part of the taxpayer in an attempt to stabilise disposable income in the face of bracket creep.**

tax changes are largely through the timing of income reporting, then a longer difference might lower the elasticity.

In the case of unlegislated tax changes such as the 'bracket creep' phenomenon investigated here, a longer window would more than likely lead to higher estimated elasticities. Over time, individual taxpayers should become more aware of the implications of 'bracket creep' and adjust their behaviour accordingly.

These issues are explored in Table 5.[29] The Table proceeds by first narrowing the window used to 2 years and then to 1 year. As is clear from the table, estimated elasticities are considerably smaller (and in one case of the wrong sign) when using a 1 year window as opposed to a 3 year window. This is consistent with the expectation that unlegislated (and relatively small) tax changes will not be immediately apparent to individual taxpayers and might take time to affect actual behaviour. Since the long-run response if of most interest, and since this is the focus of previous work, we continue to use a 3-year window in the rest of the paper.

Table 5: Variations in timing

	3-year lag	2-year lag	1-year lag
Broad income	0.139	0.061	-0.024
	[0.018]	[0.015]	[0.013]
Taxable income	0.262	0.186	0.107
	[0.019]	[0.015]	[0.014]
Observations	6 016 788	9 025 182	12 033 576

Estimates of 2SLS regressions. Regressions weighted by income. All regressions include log income, polynomials in base year income and dummy variables for each base year.

3.6 Heterogeneity

An important feature of the South African tax system, and indeed any progressive tax system, is that taxes are not linear and do not apply equally to all population subgroups. Effective tax rates differ both across income groups and between groups such as itemisers and non-itemisers. In accordance with the international literature, we have already seen that the behavioural response to a change in tax rates is much larger for itemisers relative to non-itemisers. We now proceed to consider heterogeneity across income groups.

There is significant reason to believe that the responsiveness of taxable income to changing tax rates might be higher for higher income groups, since a large share of their income comes in forms that are more readily manipulable for tax purposes. For lower income groups, the largest share of income is labour income, which is withheld for tax purposes through the pay as you earn (PAYE) mechanism. Therefore, the only way for individuals in lower income groups to manipulate income is to work more or less. For higher income households, capital and investment income will feature more prominently, and this is more readily manipulable through, for example, asset allocation decisions (Gruber and Saez, 2002).

Most papers in the international literature focus on one specific reform to identify the ETI. In these cases, most of the variation comes across income groups, making it difficult to identify group-specific responses. In contrast, Gruber and Saez (2002) use several reforms over multiple decades. These reforms impacted different points in the income distribution at different times, making it feasible to identify group-specific effects.

[29] In Table 5 and all subsequent tables we use our preferred specification that includes log income and polynomial terms in base year income as controls, and base year and itemiser dummies.

This paper (and the paper by Saez (2003)) falls somewhere in between. In South Africa, and indeed in the US, 'bracket creep' affected different points in the income distribution in different ways and at different times because nominal tax brackets were not uniformly adjusted between tax years. For example, between 2008 and 2011, the bottom tax bracket was adjusted by a cumulative 30%. However, between 2009 and 2012, the cumulative adjustment was only 17%. At the top end of the distribution, the cumulative adjustments were 18.5% and 17.5% respectively. This generates *some* within-group variability which allows us to investigate group specific behavioural responses to changing tax rates.

However, as the within-group variation due to 'bracket creep' is rather limited, the estimates are not very precise. Additionally, the lack of significant within-group variation precludes the estimation of group-specific behavioural elasticities at a fully disaggregated level. As such, we have restricted our attention to relatively broad income categories.

Table 6 gives the estimation results for these broad income categories. Columns (1) and (2) in Table 6 looks at the bottom 50%, top 50% and top 10% of income earners for broad and taxable income respectively. Income cuts are based on base year income. Column (3) provides estimates based on income tax brackets, comparing individuals in the bottom 2 tax brackets (based on base year income) to those in the middle 2 and top 2 income tax brackets.

Table 6: Elasticity estimates by income groupings

	Broad income (1)	Taxable income (2)	Taxable income (using taxable income cuts) (3)
Bottom 50% (bottom two tax brackets col. (3))	-0.118	0.0788	-0.0433
Standard deviation	[0.016]	[0.018]	[0.014]
Observations	3 008 400	3 008 387	4 113 977
Top 50% (middle two tax brackets col. (3))	0.103	0.111	-0.044
Standard deviation	[0.055]	[0.063]	[0.014]
Observations	3 008 388	3 008 401	1 287 080
Top 10% (top two tax brackets col. (3))	0.551	0.370	0.301
Standard deviation	[0.295]	[0.424]	[0.407]
Observations	601 680	601 678	615 731

Estimates of 2SLS regressions. Regressions weighted by income. All income ranges based on base year income.
All regressions include log income, polynomials in base year income and dummy variables for each base year.

It is clear from the table that most of the action in terms of the behavioural response to changing tax rates is concentrated in the higher income groups, no matter the definition used. While estimates for lower income groups are small, statistically insignificant and often of the wrong sign, the elasticity estimates for higher income groups are significantly larger. This is particularly true when looking at the top 10% of earners (for both broad and taxable income) and for individuals in the top 2 tax brackets (defined based on base year taxable income).

While the estimated elasticities for the highest income taxpayers are much larger than for lower income taxpayers, elasticities are imprecisely estimated with large standard errors. Given this imprecision, the patterns can only be taken as suggestive. Additionally, the estimated elasticities likely underestimates the response of taxpayers to changing tax rates. In fact, van Heerden (2013) estimates implied personal income tax elasticities using macro data and finds elasticities of between 0.38 and 0.79 for the various income groups under consideration.[30] Despite these caveats, the findings do confirm the standard intuition that higher income taxpayers are the most responsive to taxation.

[30] Due to lack of access to microeconomic data, the author uses macro time series to calculate the revenue-maximising and growth-maximising tax rates, which are then used to calculate the taxable income elasticity according to the Scully model (Scully, 1991).

4 Optimal Taxation

In this section, we draw on the empirical framework above to provide a computation of the optimal income tax system that is both theoretically sound and empirically based. The framework for the basic model is described in Saez (2001, 2004) and Saez et al. (2012), and summarised in Giertz (2009) and Diamond and Saez (2011).[31]

Recall that the literature on behavioural responses to taxation attempts to estimate the ETI as defined in equation (4), repeated here for convenience.

$$e = \frac{(1-\tau)}{z}\frac{\partial z}{\partial(1-\tau)}$$

As discussed in the previous section, a number of empirical studies, including this one, have found that the behavioural response to changing taxation is concentrated in in the top of the income distribution. As such, the analysis of optimal taxation will focus on higher income earners.

Following Saez et al. (2012), let's assume that individuals in the top bracket, who earn above a given reported income threshold $\bar{z}$, face a constant marginal tax rate τ.[32] The number of taxpayers in the top bracket is denoted by N. It is further assumed that the average income reported in the top bracket, z^m, depends only on the marginal net-of-tax rate, $(1-\tau)$. The aggregate elasticity of income in the top bracket with respect to the net-of-tax rate is then defined as $e = \frac{(1-\tau)}{z^m}\frac{\partial z^m}{\partial(1-\tau)}$. This aggregate elasticity is equal to the average individual elasticities weighted by individual income.[33] Most empirical studies weight individuals by their income so that they contribute to the aggregate elasticity in proportion to their incomes.

An increase in the marginal rate faced by those in the top bracket, τ, will have three distinct effects (Diamond and Saez, 2011; Kiss, 2013). The first is a mechanical effect (dM) and shows the increase in tax revenue due to the higher tax rate:

$$dM \equiv N[z^m - \bar{z}]d\tau > 0 \tag{14}$$

Second, the increase in the tax rate triggers a behavioural response that reduces the average reported income among high-income earners by $dz = -e \cdot z^m \cdot d\tau/(1-\tau)$. A change dz changes revenue by τdz. Therefore, the change in tax revenue due to the behavioural response (dB) is equal to:

$$dB \equiv -N \cdot e \cdot z^m \cdot \frac{\tau}{(1-\tau)}d\tau < 0 \tag{15}$$

These two effects combined will give the total revenue effect of any given change in the marginal tax rate τ and can be used to calculate the revenue maximising top marginal tax rate.

However, to calculate the *optimal* marginal tax rate for high income earners, we have to take the welfare loss of those who have to pay more taxes into account. This is given by:

$$dW \equiv -g \cdot N[z^m - \bar{z}]d\tau < 0 \tag{16}$$

where g is the average marginal value of consumption for those earning more than $\bar{z}$.

[31] For a more detailed discussion on optimal taxation theory, see Piketty and Saez (2012).

[32] For example, in the case of the 2008/09 tax year, individuals who earned above $\bar{z} = \text{R}490,001$ were taxed at the top marginal tax rate of $\tau = 40\%$.

[33] That is, $z^m = [z_1 + ... + z_N]/N$ and hence $e = [(1-\tau)/z^m]\partial z^m/\partial(1-\tau) = (1-\tau)[\partial z_1/\partial(1-\tau) + ...\partial z_N/\partial(1-\tau)]/Nz^m = [e_1 \cdot z_1 + ... + e_N \cdot z_N]/[z_1 + ... + z_N]$ where e_i is the elasticity of individual i.

By summing the three effects, we obtain the change in total welfare due to the increase in the marginal rate:

$$dT = dM + dW + dB = Nd\tau(z^m - \bar{z}) \cdot \Big[1 - g - e \cdot \frac{z^m}{z^m - \bar{z}} \cdot \frac{\tau}{1-\tau}\Big] \tag{17}$$

Equation 17 can be simplified and used to calculate the optimal marginal tax rate for those earning more than $\bar{z}$. The ratio $\frac{z^m}{z^m-\bar{z}}$ can be denoted by a, the so-called Pareto paremeter. The parameter a measures the thinness of the top tail of the income distribution: the thicker the top tail of the income distribution, the larger is z^m relative to $\bar{z}$ and the smaller is a. We will return to this parameter below. Using this definition of a, equation (17) can be written as:

$$dT = dM\Big[1 - g - \frac{\tau}{1-\tau} \cdot e \cdot a\Big] \tag{18}$$

At the optimal tax rate τ^*, $dT = 0$: If the effect of a small tax increase were positive (negative), the government would want to increase (cut) the tax rate further; thus the initial tax rate could not have been optimal. Setting (18) equal to zero and rearranging, we obtain the following simple expression for the optimal rate τ^*:

$$\tau^* = \frac{1-g}{1-g+a \cdot e} \tag{19}$$

The parameter g is a function of society's preferences and as such it is difficult to pin down plausible values. An upper bound to the top marginal tax rate can be obtained when setting $g = 0$. In this case, the value that society attaches to every additional rand retained by a top earner is negligible relative to an additional rand kept by the average earner (or an additional rand of government revenue). The top marginal tax rate is then given by:

$$\tau^* = \frac{1}{1+a \cdot e} \tag{20}$$

In this case, τ^* is equal to the revenue-maximising rate (Saez et al., 2012). It takes into account only the impact on total tax revenue of the mechanical (dM) and behavioural (dB) effects of a change in tax rates, ignoring the change in welfare experienced by those affected by the change.

Brewer et al. (2010) and Diamond and Saez (2011) argue that $g = 0$ is plausible for top earners (the top 1% in their definition). In fact, according to Kiss (2013), for most social welfare functions it will be the case that g is decreasing in income, and that the zero-marginal-weight result will hold asymptotically for social welfare functions that satisfy the property $\lim(g) = 0$ as z goes to infinity. Note that g does not have to converge to zero for the revenue-maximising rate to be approximately optimal. The parameter g enters (with the same sign) in both the denominator and numerator of (19) and therefore will have only second order effects as long as g is not too large. A sensitivity analysis with respect to the parameter g is performed below.

4.1 Revenue-maximising Constant Rate

Before discussing optimal tax results, it is worth considering a more straight-forward application of the estimated elasticities: the revenue maximising constant (or flat) tax rate.

Recall that $a = \frac{z^m}{z^m-\bar{z}}$ in equation (20). Note that $a = 1$ when a single flat tax rate applies to all incomes, i.e. $\bar{z} = 0$. In this case, the revenue-maximising rate (20) becomes the well-known expression $\tau^* = 1/(1+e)$. Using the results in Table 4 columns (4) and (8), we obtain a revenue-maximising rate equal to 79% for taxable income (elasticity of 0.262) and 87% for broad income (elasticity of 0.139).

As mentioned before, the elasticity estimates presented here likely underestimates the behavioural response of taxpayers to changing tax rates. In fact, using an average (unweighted) elasticity of taxable income equal to 0.65 **(calcualted as the average of the elasticity estimates for the different income groupings in van Heerden (2013))**, the revenue-maximising constant rate for taxable income falls to 61%.

Given that $a \geq 1$, the flat revenue-maximising rate is always larger than the revenue-maximising rate for top income earners only. That is because increasing the top tax rate collects revenue only on those earning in excess of $\bar{z}$, but produces a behavioural response almost as large as an across-the-board increase in the marginal tax rate (Saez et al., 2012).

4.2 Optimal Tax Results

Most studies on optimal taxation focus on the top 1% of income earners, or at least those in the top income tax bracket. However, as mentioned in Section 3.3 and discussed in Section 3.6, the value for our chosen instrument is zero in the top income tax bracket. Given that the top 1% of income earners find themselves in the top income tax bracket, this lack of within-group variation precludes the estimation of a group-specific behavioural elasticity for those individuals at the very top of the income distribution.

In what follows we will use the results from Table 6 column (2) for the top 10% of income earners to derive both the revenue-maximising and optimal marginal tax rate for those individuals at the top of the income distribution. As this grouping does not conform to an official category within the tax code, the derived optimal tax rates are purely illustrative, serving as an indication of the average optimal marginal tax rate applicable to the top 2 income tax brackets.[34]

First, the empirical value of the Pareto parameter a in equation (19) is estimated. Recall that parameter a is defined as $a = \frac{z^m}{z^m - \bar{z}}$, that is for income limit $\bar{z}$, a is equal to the average income of individuals above the limit divided by the difference of that average and the income limit. **For a Pareto distribution, the quantity $\frac{z^m}{\bar{z}}$ is constant and equal to $\frac{a}{a-1}$. Distributions with a constant $\frac{z^m}{\bar{z}}$ are exactly Pareto distributions. The tails of empirical earnings distributions can be remarkably well approximated by Pareto distributions.**

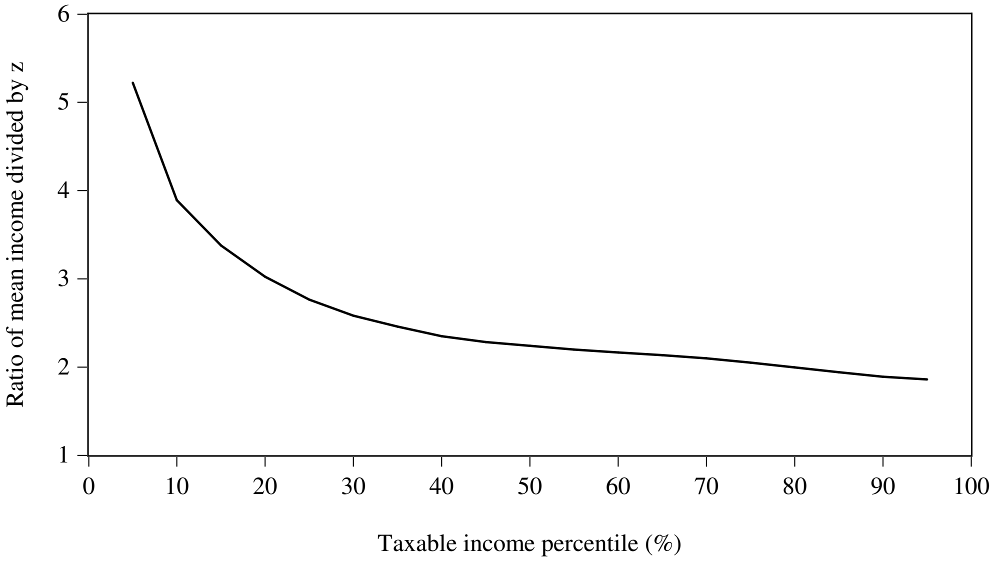
Figure 2 shows the value of $\frac{z^m}{\bar{z}} = \frac{a}{a-1}$ for the different percentiles of the taxable income distribution in our panel. The parameter declines notably between the 5^{th} and 90^{th} percentiles before stabilising at a value of just below 2, suggesting that the right tail of the income distribution is approximately Pareto distributed.

As mentioned before, the higher the parameter a, the 'thinner' the top end of the distribution. This implies that for a given value of g and e in (19), a higher value for a would result in a larger behavioural response to a change in tax rates, a larger excess burden of taxation and a lower optimal top tax rate (Steenekamp, 2012).

For the top 10% of income earners in our sample (based on real base year taxable income), $\bar{z}$ is equal to around R382 000 and z^m is equal to around R721 000, resulting in an estimate for a of 2.13. While we focus on the top 10% of income earners, this estimate compares favourably to others in the literature. **Using the same methodology and utilising data from the 2011 Tax Statistics publication**, Steenekamp (2012) estimates a value for a of 2.11 for the top 1% of South African taxpayers. This compares with estimates for the top 1% of income earners of 2.5 for Hungary (Kiss, 2013), 1.8 for the UK (Brewer and Browne, 2009), 1.6 for New Zealand (New Zealand Treasury, 2009) and 1.5 for the US (Saez et al., 2012).

[34]Across all the years in the sample, the top 2 income tax brackets contained all of the top 10% of income earners.

Figure 2: Value of $\frac{z^m}{\bar{z}}$ as a function of the chosen income limit



Revenue-maximising marginal rate

By setting $g = 0$ we can derive the revenue-maximising marginal rate for the top 10% of income earners using equation (19). From Table 4, we have that the ETI for the top 10% of income earners is 0.37. Plugging this into (19) and setting $g = 0$ and $a = 2.13$ we have that the revenue-maximising rate for the top 10% of income earners is equal to 56%. Note that this is lower than the revenue-maximising constant rate applicable to the entire income distribution as calculated in Section 4.1.

Optimal marginal rate

To calculate the *optimal* marginal tax rate, we need to take the welfare loss associated with an increase in the marginal rate, i.e. g, into account. A popular benchmark for calculating the value of g, also used by Diamond and Saez (2011), supposes that the average marginal value of consumption is inversely proportional to income. We follow Kiss (2013) in setting g equal to the inverse of the ratio of the average income of the respective income group to overall average income. For the top 10% of income earners in our sample this implies that $g = 0.28$, similar to the value estimated for Hungary in Kiss (2013).

Setting $g = 0.28$, $a = 2.13$ and $e = 0.37$ in (19) we have that the optimal marginal tax rate for the top 10% of income earners in our sample is equal to 48%.[35]

At first glance, the calculated optimal rate seems to be at odds with the legislated marginal tax rates for top income earners: over the sample period in question, legislated marginal tax rates were set at 38% and 40% for the top 2 income tax brackets respectively.

However, the actual tax rate that should be compared to the theoretical benchmark is not simply the legislated personal income tax rate. According to Kiss (2013), the tax rate corresponding to the theoretical benchmark should answer the following question: By how much can an individual increase their consumption if total labour cost is increased by one unit? Therefore, at a minimum, consumption taxes should also be taken into consideration.

The actual marginal rate that should be compared to the theoretical benchmark is then given by:

$$\tau = 1 - (1 - \tau_{cons}) \cdot (1 - \tau_{pit}) \tag{21}$$

[35] **It is important to note that, given the fact that the elasticity of 0.37 likely underestimates the behavioural response of taxpayers at the top end of the distribution to changing tax rates, the actual optimal marginal tax rate for top income earners might in fact be lower than 48%.**

where τ_{cons} is the effective consumption tax rate and τ_{pit} is the applicable marginal tax rate.

We follow Kiss (2013) and Brewer et al. (2010) in calculating the effective consumption tax rate as government revenue from domestic taxes on goods and services divided by total private consumption from National Accounts data. This results in an average effective consumption tax rate of 14.2% over the sample period in question, marginally higher than the legislated VAT rate of 14%.[36]

Using (21), setting $\tau_{cons} = 14.2\%$ and τ_{pit} equal to 38% and 40%, we have that the corresponding actual effective marginal rates for the top two income tax brackets are 46.8% and 48.5% respectively. **This corresponds well to the VAT-inclusive effective marginal rate of 47.4% as calculated in van Heerden (2013) and Steenekamp (2012) for the top income tax bracket.** This implies that the actual effective marginal tax rates for the top 2 income tax brackets (which encapsulates the top 10% of income earners) in our sample is very close to the theoretically implied optimal level of 48%. **Broadening the sample to include the 2013/14 to 2016/17 tax years and taking into account the increase in the legislated marginal tax rate for the top 2 income tax brackets to 39% and 41% in 2015, implies that the actual effective marginal rates are 47.9% and 49.7% respectively, with the latter above the theoretically implied optimal. While tax and consumption data for 2017 are not yet available, it is safe to assume that the creation of a new top income tax bracket in the February 2017 Budget (with marginal rate of 45%) implies that the effective marginal tax rate for top income earners is likely in excess of 50%, substantially higher than the calculated optimal rate of 48%.**

These calculations imply that the introduction of the new top marginal tax bracket might not yield the desired revenue results. This is confirmed by Jordaan and Schoeman (2015). The authors use a microsimulation model and suggests that the most favourable scenario for reaching the optimal PIT-to-GDP ratio would be to *lower* marginal rates. The corresponding loss in revenue will have to be compensated for, but basic theory explains that, in the longer term, the revenue base will be broadened as a result of the efficiency gains that exceed the loss in revenue.

It is important to note that effective marginal rates calculated here still do not fully reflect the actual marginal rate faced by the taxpayer as several taxes are not included in the calculation (e.g. capital gains tax, dividends tax, taxes on fringe benefits etc.). However, by including consumption taxes we do arrive at a measure that is more closely related to the optimal tax rate in equation (19).

Sensitivity analysis

As mentioned in Section 3.6, the lack of significant within-group variation resulted in imprecise elasticity estimates for the top 10% of income earners. In order to test the sensitivity of the optimal tax rate calculation, Table 7 shows the calculated optimal marginal tax rate as a function of g and e with a fixed at a value of 2.13.

Small changes in g have little effect on the optimal rate. To take an example, choosing $g = 0.1$ instead of $g = 0$ affects the optimal rate very little: for our central parameter estimates of $a = 2.13$ and $e \approx 0.4$ the optimal rate becomes 51% instead of 54%. The optimal rate is, however, substantially affected by the elasticity estimate e. The is particularly true at lower values for e: moving from $e = 0.2$ to $e = 0.3$ results in an average decrease in the optimal rate of around 9 percentage points, depending on the value of g. This moderates to a decline of around 4 percentage points when moving from $e = 0.5$ to $e = 0.6$.

Our central parameter estimates ($g = 0.28$, $e = 0.37$ and $a = 2.13$) implies that the optimal marginal tax rate for the top 10% of income earners is around 48% (consistent with a legislated marginal PIT rate of around 40%). However, as mentioned throughout the text, elasticity estimates based

[36] See the Appendix for further information on the calculation of the effective consumption tax rate.

Table 7: Optimal marginal tax rates for the top 10% as a function of g and e ($a = 2.13$)

	Values of e				
Values of g	0.2	0.3	0.4	0.5	0.6
0.0	70%	61%	54%	48%	44%
0.1	68%	59%	51%	46%	41%
0.2	65%	56%	48%	43%	39%
0.3	62%	52%	45%	40%	35%
0.4	59%	48%	41%	36%	32%

Author's calculations.

on the 'bracket creep' phenomenon likely underestimates the behavioural response of high-income earners to tax changes (see also van Heerden, 2013 for alternative elasticity estimates). Therefore, given our estimates for g and a, the optimal marginal tax rate for higher-income earners might be closer to 40%, or even below.

5 Conclusion

The magnitude of the behavioural response of taxpayers to changing tax rates is of critical importance in the formulation of tax and transfer policy, as well as for the study of the welfare implications of tax decisions.

In the absence of any large tax reforms over the sample period in question, we use the phenomenon of 'bracket creep' to estimate the elasticity of taxable income with response to changing tax rates in South Africa. The elasticity for taxable income is estimated at around 0.3, while that for broad income is estimated at closer to 0.2. These estimates lie close to the mid-point of the post-Feldstein literature. Additionally, it was found that behavioural responses are concentrated in higher-income groups as suggested by the higher elasticity estimate (0.37 for taxable income) for the top 10% of income earners.

Using these elasticity estimates as an input in an optimal taxation framework, it was found that the optimal tax rate for the top 10% of income earners is in line with a legislated marginal PIT tax rate of around 40%. Significant increases in the legislated marginal tax rate could trigger behavioural responses that would nullify any potential revenue gain. **This result is particularly important given the recent introduction of a new top income tax bracket with legislated marginal tax rate of 45%. In light of the calculations presented in this paper, the introduction of this new tax bracket might not yield the desired revenue results.**

While the use of 'bracket creep' to identify the ETI is useful in the case where large tax reforms are unavailable, several caveats do apply. First, precisely because 'bracket creep' is not a legislated tax change, taxpayers may not be fully aware of the marginal tax increases and thus might not respond to the change. This implies that one should be careful in using the estimates in the present study to predict the effects of a legislated tax reform. That being said, by using a 3-period difference in the estimation exercise, this shortcoming might be mitigated somewhat as taxpayers become more aware of the cumulative impact of 'bracket creep' on their tax liability over time and respond accordingly.

Second, we did not investigate the anatomy of the behavioural response, i.e. how different income sources respond to tax rates, which could provide insight into which sources are responsible for the behavioural response at the level of taxable income. This is arguably more important when it comes to the formulation of tax policy.

Third, the heterogeneous elasticity estimates are imprecisely measured and as such the optimal tax results should be seen as purely indicative.

Despite these caveats, the present study represents the first of its kind in South Africa and provides a solid starting point for further investigation into the behavioural response of taxpayers to changing tax rates.

References

Angrist, J. D. and J.-S. Pischke (2010). The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics. *Journal of Economic Perspectives 24*(2), 3–30.

Atkinson, A. B. and A. Leigh (2008). Top Incomes In New Zealand 1921-2005: Understanding The Effects Of Marginal Tax Rates, Migration Threat, And The Macroeconomy. *Review of Income and Wealth 54*(2), 149–165.

Auten, G. and R. Carroll (1999). The Effect Of Income Taxes On Household Income. *The Review of Economics and Statistics 81*(4), 681–693.

Auten, G., R. Carroll, and G. Gee (2008). The 2001 and 2003 Tax Rate Reductions: An Overview and Estimate of the Taxable Income Response. *National Tax Journal 61*(3), 345–64.

Blundell, R. (2006). Earned income tax credit policies: Impact and optimality: The Adam Smith Lecture, 2005. *Labour Economics 13*(4), 423–443.

Blundell, R. and T. Macurdy (1999). Labor supply: A review of alternative approaches. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, Chapter 27, pp. 1559–1695. Elsevier.

Brewer, M. and J. Browne (2009). Can more revenue be raised by increasing income tax rates for the very rich? IFS Briefing Note BN84, Institute for Fiscal Studies.

Brewer, M., A. Duncan, A. Shephard, and M. J. Suarez (2006). Did working families' tax credit work? The impact of in-work support on labour supply in Great Britain. *Labour Economics 13*(6), 699–720.

Brewer, M., E. Saez, and A. Shephard (2010). Means-testing and tax rates on earnings. In S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba (Eds.), *Dimensions of tax Design: The Mirrlees Review*, Volume 1, Chapter 2, pp. 90–201. Oxford, UK: Oxford University Press.

Burman, L. E., K. A. Clausing, and J. F. O'Hare (1994). Tax Reform and Realizations of Capital Gains in 1986. *National Tax Journal 47*(1), 1–18.

Carroll, R. and W. Hrung (2005). What Does the Taxable Income Elasticity Say About Dynamic Responses to Tax Changes? *American Economic Review 95*(2), 426–431.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics 126*(2), 749–804.

Creedy, J. (2009). The Elasticity of Taxable Income: An Introduction. Department of Economics Working Papers Series Nr. 1085, The University of Melbourne.

Diamond, P. A. and E. Saez (2011). The Case for a Progressive Tax: From Basic Research to Policy Recommendations. Technical report.

Dilnot, A. and M. Kell (1988). Top-rate tax cuts and incentives: Some empirical evidence. *Fiscal Studies 9*(4), 70–92.

Fedderke, J. (2012). The Cost of Rigidity: The Case of the South African Labor Market. *Comparative Economic Studies 54*(4), 809–842.

Feenberg, D. R. and J. M. Poterba (1993, December). Income Inequality and the Incomes of Very High-Income Taxpayers: Evidence from Tax Returns. In *Tax Policy and the Economy, Volume 7*, NBER Chapters, pp. 145–177. National Bureau of Economic Research.

Feldman, N. E. and P. Katuščák (2009). Effects of predictable tax liability variation on household labor income. *CERGE-EI Working Paper Series* (454).

Feldstein, M. (1995). The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act. *Journal of Political Economy 103*(3), 551–72.

Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics 81*(4), 674–680.

Giertz, S. H. (2007). The Elasticity of Taxable Income over the 1980s and 1990s. *National Tax Journal 60*(4), 743–68.

Giertz, S. H. (2009). The elasticity of taxable income: Influences on economic efficiency and tax revenues, and implications for tax policy. In A. D. Viard (Ed.), *Tax Policy Lessons from the 2000s*, Chapter 5, pp. 101–136. Washington, DC: AEI Press.

Gorodnichenko, Y., J. Martinez-Vazquez, and K. S. Peter (2009). Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia. *Journal of Political Economy 117*(3), 504–554.

Gottfried, P. and H. Schellhorn (2004). Empirical Evidence on the Effects of Marginal Tax Rates on Income - The German Case. IAW Discussion Papers 15, Institut fÃijr Angewandte Wirtschaftsforschung (IAW).

Gruber, J. and E. Saez (2002). The elasticity of taxable income: evidence and implications. *Journal of public Economics 84*(1), 1–32.

Hansen, J. and X. Liu (2011). Estimating Labor Supply Responses and Welfare Participation: Using a Natural Experiment to Validate a Structural Labor Supply Model. IZA Discussion Papers 5718, Institute for the Study of Labor (IZA).

Hansson, Å. (2007, October). Taxpayers' responsiveness to tax rate changes and implications for the cost of taxation in Sweden. *International Tax and Public Finance 14*(5), 563–582.

Imbens, G. W. (2010). Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature 48*(2), 399–423.

Jordaan, Y. and N. Schoeman (2015). Measuring the Impact of Marginal Tax Rate Reform on the Revenue Base of South Africa Using a Microsimulation Tax Model. *South African Journal of Economic and Management Sciences 18*(3), 380–94.

Katz Commission (1994). Report of the commission of enquiry into certain aspects of the tax structure of south africa. Pretoria: Government Printer.

Keane, M. and R. Rogerson (2012). Micro and macro labor supply elasticities: A reassessment of conventional wisdom. *Journal of Economic Literature 50*(2), 464–76.

Keane, M. P. and K. I. Wolpin (2007). Exploring The Usefulness Of A Nonrandom Holdout Sample For Model Validation: Welfare Effects On Female Behavior. *International Economic Review 48*(4), 1351–1378.

Kiss, A. (2013). The optimal top marginal tax rate: Application to Hungary. *European Journal of Government and Economics 2*(2), 100–118.

Kleven, H. J. and E. A. Schultz (2014). Estimating taxable income responses using danish tax reforms. *American Economic Journal: Economic Policy 6*(4), 271–301.

Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics 89*(11-12), 2093–2119.

Kopczuk, W., J. Friedman, H. Galper, M. Manacorda, E. Saez, D. Silverman, and J. Slemrod (2012). The polish business "flat" tax and its effect on reported incomes: a pareto improving tax reform? Columbia university working paper.

LaLonde, R. J. (1986). Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *American Economic Review 76*(4), 604–620.

Lindsey, L. B. (1987). Individual taxpayer response to tax cuts: 1982-1984: with implications for the revenue maximizing tax rate. *Journal of Public Economics 33*(2), 173–206.

Looney, A. and M. Singhal (2006). The Effect of Anticipated Tax Changes on Intertemporal Labor Supply and the Realization of Taxable Income. NBER Working Papers 12417.

Moffitt, R. A. and M. Wilhelm (2000). Taxation and the Labor Supply - Decisions of the Affluent. Technical report.

New Zealand Treasury (2009). Estimating the distortionary costs of income taxation in New Zealand. Background paper for Session 5 of the Victoria University of Wellington Tax Working Group. Available at: http://www.victoria.ac.nz/sacl/centres-and-institutes/cagtr/twg/publications/5-estimating-the-distortionary-costs-of-income-taxation-in-newzealand-treasury.pdf.

Oberholzer, R. and E. Stack (2014). Perceptions of taxation: A comparative study of different population groups in South Africa. *Public Relations Review 40*(2), 226–239.

Piketty, T. (1999). Les hauts revenus face aux modifications des taux marginaux supérieurs de l'impôt sur le revenu en France, 1970-1996. *Économie et Prévision 138*(2), 25–60.

Piketty, T. and E. Saez (2012). Optimal Labor Income Taxation. NBER Working Papers 18521, National Bureau of Economic Research.

Pirttilä, J. and H. Selin (2006). How Successful is the Dual Income Tax? Evidence from the Finnish Tax Reform of 1993. CESifo Working Paper Series 1875, CESifo Group Munich.

Robinson, Z. and R. Gcabo (2007). Tax compliance and behavioural response in South Africa: an alternative investigation. *South African Journal of Economic and Management Sciences 10*(3), 357–370.

Ross, A. M. and R. W. McGee (2012). A Demographic Study of South African Attitudes on Tax Evasion. *Journal of Economics and Economic Education Research 13*(3), 15–58.

Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies 68*(1), 205–29.

Saez, E. (2003). The effect of marginal tax rates on income: a panel study of 'bracket creep'. *Journal of Public Economics 87*(5-6), 1231–1258.

Saez, E. (2004). Reported Incomes and Marginal Tax Rates, 1960-2000: Evidence and Policy Implications. NBER Working Papers 10273, National Bureau of Economic Research.

Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy 2*(3), 180–212.

Saez, E., J. Slemrod, and S. H. Giertz (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature 50*(1), 3–50.

Saez, E., J. B. Slemrod, and S. H. Giertz (2009). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. NBER Working Papers 15012, National Bureau of Economic Research, Inc.

Saez, E. and M. R. Veall (2005). The Evolution of High Incomes in Northern America: Lessons from Canadian Evidence. *American Economic Review 95*(3), 831–849.

Scully, G. (1991). Tax rates, tax revenues and economic growth. Policy report 98, Dallas, TX: National Centre for Policy Analysis.

Selén, J. (2002). Taxable Income Responses to Tax Changes - A Panel Analysis of the 1990/91 Swedish Reform. Working Paper Series 177, Trade Union Institute for Economic Research.

Sillamaa, M.-A. and M. R. Veall (2001). The effect of marginal tax rates on taxable income: a panel study of the 1988 tax flattening in Canada. *Journal of Public Economics 80*(3), 341–356.

Singleton, P. (2011). The Effect Of Taxes On Taxable Earnings: Evidence From The 2001 And Related U.S. Federal Tax Acts. *National Tax Journal 64*(2), 323–51.

Slemrod, J. and W. Kopczuk (2002). The optimal elasticity of taxable income. *Journal of Public Economics 84*(1), 91–112.

Steenekamp, T. (2012). Taxing the rich at higher rates in south africa? *Southern African Business Review 16*(3), 1–29.

Thoresen, T. O., E. E. Bø, E. Fjærli, and E. Halvorsen (2012). A Suggestion for Evaluating the Redistributional Effects of Tax Changes: With an Application to the 2006 Norwegian Tax Reform. *Public Finance Review 40*(3), 303–338.

Thoresen, T. O. and T. E. Vattø (2013). Validation of structural labor supply model by the elasticity of taxable income. Discussion Papers Nr. 738, Statistics Norway, Research Department.

van Heerden, Y. (2013). *Personal Income Tax Reform to Secure the South African Revenue Base Using a Micro-simulation Tax Model.* Ph. D. thesis, Faculty of Economic and Management Sciences, University of Pretoria.

Weber, C. E. (2014). Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. *Journal of Public Economics 117*(C), 90–103.

Appendix

Construction of income variables

Table A.1 details the construction of the gross and taxable income variables used in the regression analysis.

Table A.1: Construction of income variables

	Item	SARS code	Notes
	Local interest	4201	Local interest earned
+	Other gains	42*	Local dividends; rental profits/losses; income from building societies; income from fixed period shares and deposits; royalties; foreign investment income (interest, dividends); gambling gains/losses
−	Other losses	42*	
=	*Investment income (excl. capital gains)*		
+	Business profits	01-34*	Profits/losses from unincorporated businesses or trades
−	Business losses	01-34*	
=	*Business income*		
+	Normal income	36*	Local and foreign labour and pension income
+	Allowances	37*	
+	Fringe benefits	38*	
+	Lump sum income	39*	Local and foreign lump-sump income, including special remuneration and pension/ provident fund lump-sums
=	*Labour income*		
=	*Gross/broad income*		All incomes received by the individual
−	Deductions	40*	E.g., pension, provident or medical fund contributions, exempted portion of interest income, other exemptions
−	Exemptions	36-39*, 42*	
=	*Taxable income*		Taxable income used to determine the normal tax due (before any rebates and tax credits)

*Asterisks refer to the subset of items under the respective SARS Code that not mentioned separately in the table.

First stage regression results

Table A.2 shows selected results for the first stage regression of the endogenous variable on the instrument (and other exogenous variables), given by

$$\log[(1-T_2')/(1-T_1')] = \log[(1-T_p')/(1-T_1')] + \theta_1 \log z_1 + \theta_2 f(taxinc_1) + \sum_{i=1}^{10} \theta_{3i} SPLINE_i(z_1) + \sum_j \theta_{4j} YEAR_j + \gamma item + \epsilon$$

with columns corresponding to those in Table 4 in the main text.

The first stage regression are very strong, suggesting instrument relevance and that we likely do not have a weak-instrument problem.

Table A.2: First stage regression results

	Taxable Income				Broad Income			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Instrument	0.612***	0.529***	0.527***	0.533***	0.610***	0.531***	0.530***	0.532***
$\left(\log[(1-T'_p)/(1-T'_1)]\right)$	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Observations	6 016 788	6 016 788	6 016 788	6 016 788	6 016 788	6 016 788	6 016 788	6 016 788
Adjusted R^2	0.074	0.108	0.117	0.153	0.073	0.107	0.114	0.139
F-statistic	99 214***	72 088***	71 084***	52 147***	98 472***	70 722***	68 894***	44 216***

*, **, *** reflect significance at the 5%, 1% and 0.1% levels respectively. Regressions weighted by income. All regressions include dummy variables for each base year.

Censored sample results

Following Gruber and Saez (2002), we censor the change in log income at 5 so that the approximately 3000 observations who report changes in income ratios across the 2 years of more than 150 or less than 1/150 are censored. This avoids the influence of large once off increases/decreases in income or job shifting. The results are not very sensitive to this restriction. Table A.1 gives the estimated elasticities (standard errors in parentheses) for both the uncensored sample and censored sample for the different sets of control variables under investigation with columns corresponding to those in Table 4 in the main text.

Table A.3: Uncensored vs censored elasticity estimates

	Taxable Income				Broad Income			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$e^c_{uncensored}$	-1.820***	0.304***	0.262***	0.314***	1.764***	0.167***	0.139***	0.235***
	[0.053]	[0.038]	[0.019]	[0.059]	[0.048]	[0.034]	[0.018]	[0.055]
$e^c_{censored}$	-1.778***	0.313***	0.271***	0.308***	1.740***	0.172***	0.144***	0.231***
	[0.053]	[0.038]	[0.018]	[0.060]	[0.048]	[0.034]	[0.018]	[0.055]

Estimates of 2SLS regressions. *, **, *** reflect significance at the 5%, 1% and 0.1% levels respectively. Regressions weighted by income. All regressions include dummy variables for each base year.

Calculating the effective consumption tax rate

Ideally, we would need to measure the tax share of the consumption basket of high-earning households. However, due to data restrictions this is not possible in practice. An approximation of the effective rate of consumption taxes is used. Data on both private consumption expenditure and government revenue from domestic taxes on goods and services were obtained from the South African Reserve Bank's (SARB) Quarterly Bulletin. Data on government tax revenue contained in the Quarterly Bulletin is in turn based on the Statement of the National Revenue, Expenditure and Borrowing as published by National Treasury.

Table A.4: Effective consumption tax rate

	2008	2009	2010	2011	2012	Average
Private consumption expenditure (current prices, Rbn)	R1 421.795	R1 492.269	R1 621.836	R1 801.091	R1 983.604	R 1 664.119
Domestic taxes on goods and services (current prices, Rbn)	R197.840	R201.141	R233.626	R266.316	R286.694	R237.123
Effective consumption tax rate	13.9%	13.5%	14.4%	14.8%	14.5%	14.2%