CENTER FOR
ECONOMIC
BEHAVIOR &
INEQUALITY

CEBI WORKING PAPER SERIES

Working Paper 23/20

IDENTIFYING BEHAVIORAL RESPONSES TO TAX REFORMS: NEW INSIGHTS AND A NEW APPROACH

Katrine Marie Jakobsen

Jakob Egholt Søgaard

ISSN 2596-44TX

CEBI
Department of Economics
University of Copenhagen
www.cebi.ku.dk

Identifying Behavioral Responses to Tax Reforms: New Insights and a New Approach*

Katrine Marie Jakobsen†

Jakob Egholt Søgaard‡

September 2020

Click here for the most recent version

Abstract

We revisit the identification of behavioral responses to tax reforms and develop a new approach that allows for graphical validation of identifying assumptions and representation of treatment effects. Considering typical tax reforms, such as a reduction in the top income tax, we show that the state-of-the-art estimation strategy relies on an assumption that trend differences in income across the income distribution remain constant in the absence of reforms. Similar to the pre-trend validation of differences-in-differences studies, this identifying assumption of constant trend differentials can be validated by comparing the evolution of income in untreated parts of the income distribution over time. We illustrate the importance of our new validation approach by studying a number of tax reforms in Denmark, and we show how violations of the identifying assumption may drive the estimates obtained from the state-of-the-art strategy (JEL: C14 H30 J22)

*We thank Amalie Jensen, Thomas Jørgensen, Henrik Kleven, Claus Thustrup Kreiner, Søren Leth-Petersen, Ben Lockwood, Sebastian Siegloch, and Floris Zoutman for helpful comments and discussions. We gratefully acknowledge support from the Danish Economic Policy Research Network (EPRN) and from the Center for Economic Behavior and Inequality (CEBI) at the University of Copenhagen, financed by grant #DNRF134 from the Danish National Research Foundation.

†University of Oxford and CEBI. Email: katrine.jakobsen@economics.ox.ac.uk

‡University of Copenhagen and CEBI. Email: jes@econ.ku.dk

1 Introduction

Behavioral responses to taxes are key inputs in evaluations of tax distortions and the trade-off between equity and efficiency (Saez, 2001; Saez *et al.*, 2012), and they serve as evidence of behavioral parameters in economic models more broadly (Chetty *et al.*, 2011a,b). Researchers and policy makers use tax reforms to identify the strength of behavioral responses to various types of taxation, including income (Feldstein, 1995; Goolsbee, 2000; Gruber & Saez, 2002; Kleven & Schultz, 2014; Weber, 2014), wealth (Seim, 2017; Jakobsen *et al.*, 2020) and corporate profits (Devereux *et al.*, 2014).

While the large literature studying behavioral responses to taxation is diverse in nature, a common challenge is that the treatment (e.g., marginal tax rate) is determined by the outcome of interest (e.g., taxable income), and thus, is endogenous. To overcome this challenge, the state-of-art estimation strategy isolates exogenous variation in treatment using tax reforms and assigns treatment status based only on pre-reform information (Gruber & Saez, 2002; Kleven & Schultz, 2014; Weber, 2014). However, as this treatment assignment is a function of past outcomes, almost any serial correlation in outcomes will violate the common trend assumption underlying differences-in-differences (DiD) studies, and serial correlation is a first order issue in the empirical tax literature as outcomes, such as wealth and income, are severely affected by secular trends and, in particular, mean reversion.

The additional challenge created by serial correlation is well known in panel data models and this is reflected in the solutions developed by the empirical tax literature. Serial correlation is dealt with either by modelling it as an AR process, which can be controlled for by including functions of past outcomes in the regressions (Gruber & Saez, 2002; Kleven & Schultz, 2014), or – inspired by Arellano & Bond (1991) – by using further lags of pre-reform information to assign treatment status (Weber, 2014). However, by relying on these solutions, the empirical tax literature has diverged from modern reduced-form studies and lacks the ability to (graphically) validate identifying assumptions. Hence, researchers often find that estimation results are highly sensitive to the exact specification (see e.g., Kopczuk, 2005) with no tools for choosing between them.

In this paper, we revisit the state-of-art estimation strategy to identify behavioral re-

sponses to tax reforms and develop a new reduced-form approach that allows for clear validation of identifying assumptions. Cast in the context of income taxation, we show that, in essence, the standard estimation strategyrelies on an assumption that any trend differences in income across the income distribution remain constant in the absence of tax reforms. Hence, we can think of the standard estimation approach as a DiDiD in time or, equivalently, a DiD with constant group specific trends. The assumption of constant trend differentials is equivalent to the common trend assumption underlying DiD studies, and we show how this assumption can be validated econometrically and graphically, in a way similar to the comparison of pre-trends in DiD studies.[1]

One way to see the DiDiD nature of the standard approach is to consider the canonical study of the US Tax Act of 1986 by Feldstein (1995). Feldstein (1995) employs a simple DiD and estimates the reform effects by comparing the change in taxable income for high and low income individuals over a period, in which the tax reform lowered marginal tax rates more for high income individuals than for low income individuals. However, with today's knowledge, this estimate is likely to be biased as the underlying income trends for high and low income individuals may be very different. To correct for this potential bias we can run the same DiD in a pre-period unaffected by tax changes and subtract the estimate from the reform DiD, thus turning it into a DiDiD. This approach yields a causal estimate of the reform effect under the assumption that the trend differences estimated in the pre-reform DiD would remain constant in the absence of the reform. In its simplest form, this is what the standard approach does by controlling for past income.[2]

Our new approach builds on the DiDiD to validate the assumption of constant trend differentials. The key insight is that individuals treated similarly by a tax reform (e.g., individuals within a given tax bracket) are not a homogeneous group, but are drawn from a wider income range. Thus, within these groups we would expect differences in income trends, but, under the assumption of constant trend differentials, we should observe no

[1]One might note that a DiDiD yields one estimate of the reform effect and does not immediately explain the sensitivity of the standard approach to the exact specification. However, in more elaborate specifications the assumption of constant trend differentials is conditional on the additional controls included, and hence, different sets of controls will fulfill or violate the assumption of constant trend differentials to differing extents.

[2]We also show how this insight extends to the use of further lags of pre-reform information to assign treatment status, as done by Weber (2014).

changes in these trend differentials within the untreated parts of the income distribution. Hence, we can validate the identifying assumption by non-parametrically comparing trend differentials in two periods: a reform period, where a reform changes tax rates differently for different groups, and a pre-reform period, where ideally the tax system was stable.

We illustrate our new approach using a number of tax reforms in Denmark with a particular focus on two reforms implemented in 2004 and 2009-10. The 2004 reform primarily reduced marginal tax rates in the middle of the income distribution, while the 2009-10 reform primarily reduced them at the top, with only minor changes in the bottom.

Applying our new estimation approach to the 2009-10 tax reform, we find significant negative correlations between initial income and subsequent income growth, which is consistent with mean reversion being the dominant (but not necessarily the sole) feature of the underlying income process in both the pre-reform and reform periods. Next, comparing the pre-reform period to the reform period, we find that the income trend differentials remained stable for the untreated bottom part of the income distribution, while for the treated upper part of the income distribution we find significantly higher income growth. The changes in trend differentials are strongly increasing in initial income and somewhat larger in the medium run than in the short run. Taken together, we see this as compelling evidence of behavioral responses to the 2009-10 tax reform.

The behavioral responses translate into an average elasticity of taxable income with respect to the marginal net-of-tax rate of 0.227, ranging from 0.1 in the middle of the income distribution to 0.5 at the top.[3] We find that most of the responses are driven by income shifting from pension contributions to taxable income induced by the reform. Using a broader income measure that is unaffected by the shifting of income, we find an elasticity of 0.016.

Turning to the 2004 reform, we find close to the same average elasticities as with the 2009-10 reform when we apply the standard approach. However, inspecting the trend differentials across the income distribution we find no changes around the changes in tax

[3]Our finding of heterogeneous elasticities would normally invalidate the point estimates. This is less of a case in our setting, where we have close to a pure control group untreated by the tax reform. See also Kumar & Liang (2020).

treatment. Instead, the elasticity estimate is driven by changes in trend differentials well within the control group and is most likely unrelated to the reform. Once we account for these changes, we obtain a precisely estimated zero response for both taxable and broad income.

The potential biases that we address are not new. The problems of mean reversion and secular differences in income trends were highlighted already by Auten & Carroll (1999) and discussed extensively in the large literature that followed (see e.g. Saez *et al.*, 2012). Our contribution is to bring the empirical tax literature up to modern empirical standards by clarifying the assumptions underlying the state-of-the-art estimation strategies and develop tools to (graphically) validate these assumptions. Other issues in the literature, such as shifting across tax bases (Slemrod, 1995; Gordon & Slemrod, 1998; Pirttilä & Selin, 2011), shifting across time (Goolsbee, 2000; Kreiner *et al.*, 2016), the endogeneity of the responses to the design of the tax system (Slemrod & Kopczuk, 2002; Kopczuk, 2005; Fack & Landais, 2016) and the presence of optimization frictions (Chetty, 2012), are all related to the expected size and anatomy of behavioral responses to tax reforms and naturally need to be dealt with in a comprehensive empirical analysis. However, they are orthogonal to our points on the empirical identification of responses.

Finally, while our paper focuses on the empirical tax literature, our new approach applies to studies of other policies where treatment is a direct or indirect function of the outcome variable of interest. These include, for example, minimum wage policies, where one type of policy evaluation compare changes in outcomes for the treated bottom part of the hourly wage rate distribution to changes in other parts of the distribution (see, e.g., Currie & Fallick, 1996, Clemens & Wither, 2019 and Dustmann *et al.*, 2019). Similarly, we could consider the consumption effects of drops in housing prices for highly indebted individuals (Mian & Sufi, 2010) or the effect of size-dependent firm regulation (Garicano *et al.*, 2016).

The rest of the paper is organized as follows. Section 2 revisits the state-of-art strategy for estimation of behavioral responses to taxation and outlines our new estimation approach. Section 3 introduces the institutional setting and data used in our empirical application, while Section 4 presents graphical evidence on income responses to the 2004 and

2009-10 reforms. Finally, Section 5 concludes.

2 Estimating Behavioral Responses to Taxation

In this section, we develop our new approach for estimation of behavioral responses to taxation. We will do so in three steps. First, we set up a simple theoretical framework in the context of income taxation and show the basic difficulties in estimating income responses to different types of tax variation. Second, we revisit and reinterpret the state-of-the-art estimation strategies starting from the work of Gruber & Saez (2002).[4] Finally, we layout a new estimation approach that allows for graphical validation and identification of income responses to tax reforms.

Theoretical Framework

To fix ideas we start by setting up a simple model for the "supply" of taxable income.[5] Each individual i at time t maximizes the following quasi-linear utility function

$$u_{it}(c_{it}, z_{it}) = c_{it} - \frac{n_{it}}{1 + \frac{1}{\varepsilon}} \left(\frac{z_{it}}{n_{it}} \right)^{1+\frac{1}{\varepsilon}}, \tag{1}$$

subject to a potentially non-linear budget set

$$c_{it} \leq z_{it} - T_t(z_{it}; x_{it}), \tag{2}$$

where c_{it} is consumption, z_{it} is taxable income and x_{it} is a set of other variables, such as the number of children, marital status or underlying components of taxable income that may affect individuals' tax liability $T_t(z_{it}; x_{it})$ in addition to their taxable income. The last term

[4] We start from the work of Gruber & Saez (2002) and not the more recent work of Weber (2014), as our reinterpretation of the Gruber & Saez (2002) approach is also the logical starting point for reinterpreting the Weber (2014) approach.

[5] It is custom in the elasticity of taxable income (ETI) literature to implicitly ignore the "demand" side of taxable income. In part this goes back to the idea of the ETI being a sufficient statistic for the computation of the deadweight loss (Feldstein, 1995, 1999), which rests on the assumption that the "price" of taxable income is fixed. If instead, for example, an increase in hours worked reduces the hourly wage rate, using the ETI will underestimate the deadweight loss.

of equation (1) captures the disutility associated with earning income, which is governed by the parameters n_{it}, and ε. These parameters can – as we show below – be interpreted as potential income (absent taxation) and the elasticity of taxable income with respect to the marginal net-of-tax rate.[6] As in other studies of taxable income responses, we think of z_{it} as being a function of a range of underlying margins, such as hours worked, choice between pecuniary and non-pecuniary job attributes, form and timing of compensation, tax avoidance, and evasion, all of which may be part of individuals' responses to tax changes.

Utility maximization yields the first order condition

$$z_{it} = n_{it} \left(\tau_{it}\right)^{\varepsilon}, \tag{3}$$

where $\tau_{it} = 1 - T_t'(z_{it}; x_{it})$ is the individual marginal net-of-tax rate in year t. Rewritten as log differences, equation (3) becomes

$$\begin{aligned} \ln z_{it} - \ln z_{it-1} &= \varepsilon(\ln \tau_{it} - \ln \tau_{it-1}) + \ln n_{it} - \ln n_{it-1} \\ \Leftrightarrow \triangle \ln z_{it} &= \varepsilon \triangle \ln \tau_{it} + \triangle \ln n_{it}. \end{aligned} \tag{4}$$

Econometrically, we are interested in the causal effect of a change in the marginal net-of-tax rate on taxable income captured by the elasticity ε in equation (4). However, as is evident from equation (4), a naive estimation of log taxable income on the log of the net-of-tax rate gives biased estimates as the marginal tax rate is a (deterministic) function of the dependent variable and, hence, the error term (the change in potential income, $\triangle \ln n_{it}$).

To break this endogeneity problem, researchers, starting with Auten & Carroll (1999), have employed a simulated instrumental variable (IV) strategy using tax reforms.[7] The basic idea in this approach is to use predicted changes in marginal net-of-tax rates driven only by changes in the tax system

$$\triangle \ln \tau_{it-1}^{p} = \ln\left(1 - T_t'(z_{it-1}; x_{it-1})\right) - \ln\left(1 - T_{t-1}'(z_{it-1}; x_{it-1})\right) \tag{5}$$

[6] In the main analysis, we ignore the possibility of income effects and assume a homogeneous elasticity (ε) across the population. We return to both of these points at the end of Section 2.

[7] The DiD used by Feldstein (1995) was, in practice, a Wald estimator and, hence, not fundamentally different from the IV strategy used in later research.

as an instrument for $\triangle \ln \tau_{it}$. However, this basic IV strategy only yields causal estimates of ε if the exclusion restriction holds:

$$\operatorname{cov}\left(\triangle \ln \tau^{p}_{it-1}, \triangle \ln n_{it}\right) = 0. \tag{6}$$

In other words, the instrument given by the predicted changes in marginal net-of-tax rates should be uncorrelated with the changes in potential income.

Why the Basic IV Strategy is Unlikely to Work

To think about the exclusion restriction (6), it is useful to distinguish between two types of tax variations. Defining $\triangle \ln \bar{\tau}^{p}_{zt-1} = E\left(\triangle \ln \tau^{p}_{it-1} \middle| z_{it-1}\right)$ as the average reform-driven change in the net-of-tax rate given initial income z_{it-1}, we can break up the total tax variation into *between income* and *within income* tax variation

$$\underbrace{\operatorname{var}\left(\triangle \ln \tau^{p}_{it-1}\right)}_{\text{Total Variation}} = \underbrace{\operatorname{var}\left(\triangle \ln \bar{\tau}^{p}_{zt-1}\right)}_{\text{Between Income Tax Variation}} + \underbrace{\operatorname{var}\left(\triangle \ln \tau^{p}_{it-1} - \triangle \ln \bar{\tau}^{p}_{zt-1}\right)}_{\text{Within Income Tax Variation}}. \tag{7}$$

The between income variation is created by reforms that change marginal tax rates differently for different income groups (such as a reduction in a top tax), while the within income variation is created by reform changes in marginal tax rates that differ among individuals with the same initial income level (such as tax credits for individuals with children or treatment of itemized deductions).

We distinguish between these two types of tax variation, because the empirical strategies differ in the two cases. Below we focus on between income variation, which is the stereotypical tax variation considered in the literature (Saez *et al.*, 2012; Weber, 2014), and we delegate the treatment of within income variation to the end of Section 2 and – in greater detail – to Appendix C.

With reforms that only create between income variation, the exclusion restriction (6) will most likely not hold. To see this, note that the instrument τ^{p}_{it-1} is a function of two elements: the change in the tax schedule created by the reform and the individuals' initial income, z_{it-1}, which in turn is a function of n_{it-1}. Hence, if a tax reform changes marginal

tax rates in a way that is correlated with income and if the changes in potential income $(\triangle \ln n_{it})$ are correlated with the initial level (n_{it-1}), the exclusion restriction does not hold.

Most tax reforms do change marginal tax rates in a way that is correlated with income because most reforms aim at adjusting the balance between efficiency and redistribution of the tax system and, hence, adjust the progressivity of the tax schedule. A prominent example of this is the US tax act of 1986 (TRA86) studied by, for example, Feldstein (1995), which lowered marginal tax rates much more at the top of the income distribution than at the bottom.

Similarly, there are at least two reasons why changes in potential income are correlated with the initial level:

- **Mean reversion:** if part of n_{it-1} is driven by temporary shocks, individuals with higher z_{it-1} are more likely to experience a negative change in potential income. Hence, $E\left(\triangle \ln n_{it} | z_{it-1}\right)$ is decreasing in z_{it-1}.
- **Differential income trends:** e.g., secular increases in inequality, where individuals with higher z_{it-1} on average have larger growth in potential income for reasons other than the tax reform. In this case, $E\left(\triangle \ln n_{it} | z_{it-1}\right)$ would be increasing in z_{it-1}.

These potential biases are well known in the literature and, in most studies, mean reversion appears to be a severe and dominant source of bias. When the variation in marginal tax rates predominately comes from larger reductions at the top of the income distribution, mean reversion biases the estimate downward, which is consistent with the negative elasticity estimates usually obtained in the literature when using the basic IV strategy (Gruber & Saez, 2002; Kopczuk, 2005; Kleven & Schultz, 2014; Giertz, 2015).

Reinterpreting the Gruber and Saez (2002) Approach

To deal with the likely violation of the exclusion restriction, researchers have – again dating back to Auten & Carroll (1999) – included controls for initial income (z_{it-1}) in the

estimation, which changes the exclusion restriction to

$$\text{cov}\left(\triangle \ln \tau_{it-1}^{p}, \triangle \ln n_{it} | z_{it-1}\right) = 0. \tag{8}$$

Looking only at the reform period (changes from before to after the reform) this condition is fulfilled trivially in cases with only between income variation. In these cases, $\triangle \ln \tau_{it-1}^{p} = \triangle \ln \bar{\tau}_{zt-1}^{P}$ is a constant for a given level of initial income, and equation (8) thus holds regardless of the changes in potential income. However, controlling fully non-parametricly for initial income as implied in equation (8) also absorbs all identifying variation in the tax instrument in the first stage.[8]

There are two ways to break this deadlock. We can abandon the non-parametric controls for initial income and assume a functional form of the relationship between $\triangle \ln n_{it}$ and z_{it-1}. Or we can employ more time periods to obtain variation in $\triangle \ln \tau_{it-1}^{p}$ conditionally on z_{it-1}. Auten & Carroll (1999) only had one period available, and hence, they were only able to control for initial (log) income linearly. In general subsequent studies have had access to longer panels and have thus been able to use less parametric specifications.

With longer panels, $\triangle \ln \bar{\tau}_{zt-1}^{P}$ only varies due to the time variation created by tax reforms. Hence, a sufficient condition for equation (8) to hold is that the dynamic process of potential income given initial income does not change systematically over time. Formally we can write this condition as

$$E\left(\triangle \ln n_{it} | z_{it-1}\right) = \tilde{g}(z_{it-1}) + \delta_t, \tag{9}$$

where $\tilde{g}(z_{it-1})$ is some time-independent function of initial income describing the dynamic income process, and where δ_t is a common income growth rate.[9]

[8] The first stage equation corresponding to equation (8) is $\text{cov}(\triangle \ln \tau_{it-1}^{p}, \triangle \ln \tau_{it} | z_{it-1})$, which is also zero when $\triangle \ln \tau_{it-1}^{p} = \triangle \ln \bar{\tau}_{zt-1}^{P}$.

[9] One way to think about equation (9) is to consider the following process of potential income:

$$\triangle \ln n_{it} = g_t\left(z_{it-1}, x_{it}\right).$$

This is a completely general process for potential income stating that the change in potential income can be a function of initial income and other covariates (x_{it}), such as age, experience, children etc. and that this function may change over time. Within this setting, there are two sufficient conditions for equation (9) to hold. First, $g_t\left(z_{it-1}, x_{it}\right)$ must be independent of time up to a constant: $g_t\left(z_{it-1}, x_{it}\right) = g\left(z_{it-1}, x_{it}\right) +$

Looking at equation (9), we see that the equation is an assumption of constant trend differentials across the income distribution over time in the absence of the reform. In the absence of changes in the tax schedule, changes in taxable income are solely driven by changes in potential income, and equation (9) states that, relative to the overall growth in the economy, the differences in income growth across the distribution should be constant. This is the key identifying assumption underlying the state-of-art strategy as employed by, for example, Gruber & Saez (2002) and Kleven & Schultz (2014).[10]

A New Approach: Graphical Validation and Identification of Behavioral Responses

With the above insight, the identification of income responses and the validation of the identifying assumption underlying equation (9) are straightforward. For identification, we compare the changes in income trends for the parts of the income distribution affected by tax changes to the changes in income trends for the untreated (or less treated) parts of the distribution, while for validation, we can compare the changes in income trends for different subgroups within the untreated parts of the distribution. Under the assumption of constant trend differentials, we should observe no changes in trend differentials within the untreated (or less treated) parts of the income distribution.

We illustrate our new approach in Figure 1, which shows the growth in income across the income distribution for two time periods: a pre-reform period, where the tax schedule remains stable and a reform period, where the top tax rate is reduced. In both the pre-reform and the reform period, we draw the growth in income as a decreasing function of initial income (z_{t-1}). I.e., income trends differ across the income distribution with individ-

δ_t. Second, the distribution of other covariates conditional on initial income must be constant over time: $F_t(x_{it}|z_{it-1}) = F(x_{it}|z_{it-1})$. If the distribution of x_{it} is not constant, we can control for these in the estimations as discussed at the end of Section 2.

[10] In some DiD studies researchers correct for differential pre-trends by estimating a linear trend using pre-reform data and subtracting this from the post-reform estimates. Hence, the identifying assumption in these studies is also one of constant trend differentials. However, this approach is unlikely to work in the case of a tax reform when we assign individuals to treatment and control groups based on their income in the year prior to the reform, as mean reversion creates very different income trends going forward and backward from this point. After the year of assignment, mean reversion tends to reduce the income of high income individuals relative to low income individuals, while income growth tends to be higher for the (to be) high income individuals prior to the point of assignment. Hence, only in very specific cases are we able to study tax reforms in a DiD framework. See also the discussion of within and disperse between variation at the end of Section 2.

uals at the top experiencing lower income growth, on average, than individuals with low initial income. This pattern is consistent with mean reversion being the dominant feature of the underlying income process, but it is not necessarily the only feature. Importantly, we do not impose any functional form assumptions on the underlying income process.

Next, we compare the trend differentials in the two periods, and for this purpose, we divide the income distribution into two regions. An untreated *validation region*, where the comparison of trend differentials serve the same validation role as the comparison of pre-treatment trend in DiD studies, and treated *identification region*, where we – under the identifying assumption – can interpret changes in trend differentials as behavioral responses to the reform. In our illustration, the trend differentials in the validation region follow the same pattern in the pre-reform period as in the reform period, while the trend differentials for the reform period lie above the pre-reform period. Thus, in our illustration we find would positive reform effects with clear validation of the identifying assumption.

It is worth noting that as our new approach builds on the same identifying assumption as the standard approach, it will not (necessarily) differ in terms of estimated reform effects. Indeed, if our graphical validation exercise carries through, we can use the standard approach to translate the observed reform effects into an aggregate income elasticity. However, without validation, any aggregate elasticity driven by the standard approach may just be produced by changes in trend differentials within the validation region, which is a strong indication of violations of the identifying assumption. We illustrate this issue empirically in Section 4 along with the practical implementation of our new approach.

Extensions and Additional Considerations

Using further lags of initial income: In her recent paper, Weber (2014) suggests using further lags of initial income when computing the predicted changes in marginal net-of-tax rates to reduce the mean reversion problem. Intuitively, this is likely to work as any transitory component in potential income (by definition) will play out over time, and hence, not affect changes in potential income sufficiently far into the future. In Appendix B we show that this approach can be reinterpreted as a natural extension of our approach, where we compare changes in trend differentials across the income distribution defined using fur-

ther lags of initial income (z_{it-1-k} for some $k > 0$) instead of initial income (z_{it-1}) as in Figure 1.

Adding Additional Controls: When applying the standard approach, researchers often include additional controls in the estimations with the aim of better modelling the underlying income process. These include controls such as changes in past income and more traditional controls for demographics etc. Additional controls can also be added to our new approach either directly in the estimation equations or by applying weights. In this case, the assumption of constant trend differentials needs to hold conditional on the controls, which we can validate by showing either the residualized or reweighted income trends in Figure 1.

Within Income Tax Variation: Most tax reforms create both between and within income tax variation, and while our new approach focuses on the use of between variation, we show in our empirical application how to use the two types of variation separately. With within income tax variation we do not face the same immediate problem of differential income trends due to mean reversion, as we can identify income responses by comparing individuals with similar initial income. Hence, as discussed in Appendix C, we can – with a few adaptation – analyze this type of tax variation in a DiD framework.

Disperse Between Income Tax Variation: In cases where tax reforms create between income variation in marginal tax rates that is sufficiently dispersed throughout the income distribution, the variation may come close to being uncorrelated with initial income, and hence, allow a DiD analysis similar to the case with within income tax variation. One example of this is Saez (2003), who studies the large "bracket creeps" in the US in the late 1970s. In this case, treatment and control groups were distributed throughout the income distribution instead of being concentrated at the top, as in the example above.[11]

Heterogeneous Elasticities: In our theoretical framework we made the simplifying as-

[11] In the case of Saez (2003), the changes in marginal tax rates turned out to be correlated with income. Thus, Saez (2003) still includes controls for initial income. However, the particular source of variation does enable Saez (2003) to replace the requirement of a panel with at least two periods with an assumption that the relationship between the change in income and initial income, $\tilde{g}(z_{it-1})$ in equation (9) , is sufficiently smooth. Intuitively, Saez (2003) compares a group of individuals, who are pushed into another bracket by high inflation with individuals with income just below and just above, who remain in their bracket. Thus, the counterfactual income growth of the treatment group can be estimated as a weighted average of the income growth of the two "control" groups.

sumption of a constant elasticity across the population. It is well-known that heterogeneous elasticities bias the elasticity estimates when there is no pure control group (with unchanged marginal tax rates), but only differences in treatment intensities. The same applies in our approach, but as we show in Section 4, it is straightforward to investigate whether the elasticities are heterogeneous and, hence, to determine whether they are likely to be biased.

Income Effects: Income effects to tax changes are just one of many reasons why the (uncompensated) elasticities considered in our theoretical framework could vary across the population. Hence, similar to the detection of heterogeneous elasticities, we can use our approach to investigate whether the heterogeneity of the elasticities follows a pattern predicted by the presence of income effects. We discuss these additional challenges in Appendix D.

Pooling Multiple Tax Reforms: One version of the standard estimation approach is to pool the variation created by multiple tax reforms in a single regression with the aim of obtaining average elasticity across all reforms. Similar to our above discussion of applying the standard approach to a single tax reform, the pooled approach does not necessarily lead to biased estimates. However, it is worth noting that pooling tax reforms makes it more difficult to validate the assumption of constant trend differentials, which is problematic as identification in the pooled case rests on the assumption that the underlying trend differentials remain constant over longer time periods.

3 Institutional Setting and Data

Institutional Setting

In Denmark, individuals are taxed according to a dual tax system with generally higher taxes on labor income (and transfers) than on capital income.[12] The system operates with six income concepts summarized in Table 1. The income concepts are labor market income (LI), personal income (PI), capital income exclusive of income from stocks (CI), stock

[12]This section extends the description of the Danish tax system provided by Kleven & Schultz (2014), who cover the tax reforms between 1984-2004.

income (SI), itemized deductions (ID) and taxable income (TI) with LI, PI, and TI constituting the main tax bases for labor income.[13] Overall, these income concepts have remained stable over the time period 2000-12, which surrounds the 2004 and 2009-10 reforms we consider in our main application.

Table 2 shows the tax bases and associated marginal tax rates before and after the two reforms. The tax system consists of a number of flat elements levied on the entire base with only minor allowances (labor market contribution, regional taxes, bottom tax) combined with progressive elements created by the income thresholds of the middle and top tax brackets. The system creates three overall tax brackets, which we label *bottom*, *middle,* and *top*.

The tax rates shown in the table are cumulative such that a taxpayer in the top tax bracket is subject to the sum of all tax rates (except the tax on stock income). Hence, a top taxpayer faced a marginal tax of 63.0 percent on labor income before the 2009-10 reform and 56.1 percent after. This corresponds to a change in the net-of-tax rate of 17 log points. In contrast, the marginal tax rate for a bottom taxpayer is only reduced by 3 log points from 42.6 to 40.9 percent.

Figure 2 illustrates the development in the marginal tax rates on labor income (Panel A) and income thresholds (Panel B) for the three overall tax brackets for a typical tax payer. There are a number of important points to take away from this figure. First, in the period leading up to the 2009-10 reform the income tax system was stable in terms of both marginal tax rates in each tax bracket and the bracket thresholds. Second, the 2009-10 reform affected the tax system in two ways. In 2009, the middle tax threshold was increased to be the same level as the top threshold, and in 2010 the middle tax bracket was abolished entirely.[14] Because the tax rates are cumulative, this affects all income above the (old) mid-

[13]The duality is created by only taxing capital income as taxable income (TI), while labor income, in addition, is taxed as personal (PI) and labor income (LI). Only stock income (SI) is taxed completely separately from other income concepts. Capital income is negative for the majority of Danish taxpayers due to interest payments on mortgages and other loans, while the typical sources of positive capital income (e.g. stock income) are taxed separately. On top of the direct taxation of income, Denmark also levies a 25% VAT on close to all goods and services.

[14]The law abolishing the middle tax in 2010 was passed in parliament already in May 2009, which gave individuals more than half a year to shift income from 2009 to 2010 and thus avoid the middle tax. This created significant income shifting from 2009 to 2010 among certain groups of high-income earners (see Kreiner *et al.*, 2016). We deal with these avoidance responses by primarily focusing on changes in income

dle tax threshold, including income above the top tax threshold. Third, after the 2009-10 reform, the tax system is again relatively stable until 2012 with only minor changes in the top tax threshold. Taken together, these features come close to the stylized tax reform considered in Section 2 with a stable pre-period and reform that generates substantial between income tax variation through the initial threshold increase and subsequent abolishtion of the middle tax.

In addition, it is worth noting that the 2009-10 reform introduced a cap on certain deductible pension contributions. Prior to the 2009-10 reform all contributions to employer administrated annuity pensions were deducted from personal income and hence tax exempt at the point of contribution.[15] With the reform, the government introduced a cap of DKK 100,000 ($\approx$ USD 15,000) on expiring annuities (typically with payouts over 10 years). Many employers reacted to the cap by automatically shifting contributions from expiring to non-expiring annuities, but as shown by Andersen (2018) many of the affected tax payers reduced overall pension contributions in response. In our analysis, this shifting of income will be captured as part of the behavioral responses to the 2009-10 reform, and for this reason we consider income responses both with and without pension contributions in our empirical analysis in Section 4.

Compared to the 2009-10 reform, the 2004 reform comes after a more unstable pre-period with gradual declines in the marginal tax rates in all brackets and a gradual increase in the middle tax threshold, and instead of changing statutory tax rates, the main element of the 2004 reform is a significant increase in the middle tax threshold.

Moving outside the 2000-12 window, the Danish tax system is affected by a number of other reforms. In Figure 2, we also highlight the tax reforms in 1987 and 1994, which we analyze in Appendix E.

from 2008 to 2012.

[15] Employer administrated pensions in Denmark resemble the US 401(k) accounts (Chetty *et al.*, 2014) with generally fully tax exempt contributions, while pension payouts are taxes as regular income. Hence, top tax payers could often reduce their total taxes by shifting income from high income employment periods to retirement as pensions contributions.

Data

We use administrative data for the full population of Denmark since 1980. The data combine several administrative registers (linked at the individual level via personal identification numbers) that contain detailed information on labor market history, education, earnings, and demographics with almost all income data being third-party reported (Kleven *et al.*, 2011). However, our new approach does not rely on the availability of a rich set of control variables and is, therefore, implementable using data from tax authorities only, which is increasingly becoming available to researchers and policy makers.

Crucially, individual marginal tax rates are not observed in the data. Like most other papers in the empirical tax literature, we address this challenge by imputing marginal tax rates using a detailed tax simulator (similar to the NBER TAXSIM). For this purpose, we extend and use the model of the Danish tax system created by Kleven & Schultz (2014) and compute marginal tax rates as the increase in total household taxes due to an incremental increase in labor income.[16]

For each reform, we construct a pre-reform period leading up to the reform and a reform period spanning the reform. As our baseline we consider four year periods so that, for example, the pre-reform and reform periods for the 2004 reform span 1999-2003 and 2003-07, respectively. For each period, we select individuals who are present in the tax data with positive income over the period considered and with initial income within a range around the tax variation created by the reform. For the 2009-10 reform we select individuals with personal income above DKK 250,000 (USD 37,000, 2019-level), while for the 2004 reform we include individuals with personal income starting already from DKK 200,000 (USD 29,000, 2019-level) as the main tax variation from this reform is located lower in the income distribution. However, the conclusions from our empirical analysis are not sensitive to the sample restrictions.

The sample restrictions leave us with over 2,000,000 individuals per period or approximately half of the Danish adult population, as shown in Table 3. In the table, we also show

[16] A simpler approach is to assign marginal tax rates based on the statutory rates (listed in Table 2) and individuals' personal income, which is, with minor deviations, the income concept affected by the 2009-10 reform. This approach comes close to an estimation using only the between income tax variation as it misses the within variation created by changes in credits for individuals with children, treatment of itemized deductions, etc.

the descriptive statistics for the pre-reform and reform periods for the 2004 and 2009-10 reforms, respectively. These statistics are measured in the initial year of each period. Considering the 2004 reform, we see that the initial characteristics are stable from the pre-reform period (initial year 1999) to the reform period (initial year 2003). As illustrated in Figure 2, the main feature of the 2004 reform is an increase in the middle tax threshold from a personal income of DKK 290,000 to DKK 360,000, which, as shown in Table 3, corresponds to an increase from just above the 25th percentile to just above the median. For the 2009-10 reform, we only include individuals with initial income above DKK 250,000, and hence, the samples are, on average, older, more likely to be married and have higher income. As the main feature of the 2009-10 reform is an abolition of the middle tax, the reform affects all those with personal income above the new middle tax threshold of DKK 360,000, which lies between the 25th percentile and median in the sample.

4 Graphical Evidence of Income Responses

In this section, we turn to the empirical implementation of our new approach for studying income responses to taxation using the 2004 and 2009-10 reforms described in Section 3. We start by analyzing the 2009-10 as this is the largest and most salient of the two and as this reform changes marginal tax rates in a way that comes close to the stereotypical reform considered in Section 2.

The 2009-10 Tax Reform

We implement our new approach in four steps, as illustrated in Figure 3. Starting in Panel A, we show the predicted changes in the marginal net-of-tax-rate across the income distribution for two periods.[17] First, in the pre-reform period from 2004 to 2008, where the tax schedule was close to stable, we see close to no changes in marginal net-of-tax rates both across the income distribution and, as illustrated by the P10-P90 range, within

[17] Unless stated otherwise, we use personal income (PI in Table 1) as (taxable) income since this is the relevant income concept affected by the 2009-10 reform. We generally refer to taxable income responses as changes in the income concept subject to a specific tax change. Hence, it should not be confused with the legal concept of taxable income defined in the tax system (TI in Table 1).

a given income level. Second, in the reform period spanning 2008 to 2012, the abolition of the middle tax increased the marginal net-of-tax rate by approximately 17 log points for the highest income individuals, while the average changes for individuals with incomes below DKK 350,000 are less than 5 log points.[18]

Next, in Panels B and C of Figure 3, we investigate the income trend differentials across the income distribution in the pre-reform (2004-08) and reform (2008-12) periods. To construct Panel B, we run the following regression separately for each period t

$$\Delta \ln z_{it} = \beta_0^t + \boldsymbol{\beta_1^{t\prime}} \boldsymbol{D_{it}^{inc}} + \nu_{it}, \tag{10}$$

where $\boldsymbol{D_{it}^{inc}}$ is a vector of 50 initial income bin dummies based on initial income measured in 2004 or 2008. We exclude the dummy for the DKK 300,000 income bin, and hence, the coefficients $\boldsymbol{\beta_1^t}$ measure the differences in income growth (income trend differentials) between individuals with different initial income relative to the DKK 300,000 bin. Plotting $\boldsymbol{\beta_1^t}$ for the two periods produces the empirical equivalent of our stylized Figure 1 in Section 2. We consider changes in income over the full 4 year period to allow time for potentially gradual behavioral responses to materialize and to avoid short term income shifting affecting the estimates. In particular, we avoid capturing the income shifting from 2009 to 2010 documented by Kreiner *et al.* (2016, 2017).[19]

In Panel C we estimate the changes in trend differentials between the pre-reform and reform periods by running the regression

$$\Delta \ln z_{it} = \delta_0 + \boldsymbol{\delta_1}' \boldsymbol{D_{it}^{inc}} + \delta_2 D_{it}^{reform} + \boldsymbol{\delta_3}' \boldsymbol{D_{it}^{inc}} \times D_{it}^{reform} + \nu_{it}. \tag{11}$$

In this estimation, we interact the vector of income bin dummies $\left(\boldsymbol{D_{it}^{inc}}\right)$ with a dummy for the reform period $\left(D_{it}^{reform}\right)$. Hence, the coefficients $\boldsymbol{\delta_3}$ capture the changes in income

[18] The spike around DKK 450,000 is caused by a small increase in the top tax threshold, which, together with the abolition of the middle tax, moved some individuals from the highest to the lowest tax bracket. The 2009-10 reform also created within income tax variation due to the fact that the middle tax was based on couples' total income. Hence, differences in spousal income created differences in the income level at which individuals start paying the middle tax rate. We study this source of variation in Appendix C.

[19] In Appendix Figure A.I. we consider income changes over 2 and 3 years to show how the 4-year effects on taxable income build up over time.

trend differentials across the income distribution. We estimate the changes in trend differentials for different groups and income concepts: *Taxable income*, which corresponds to the estimates from Panel B; *Excluding self-employed,* which excludes all individuals with income from self-employment; and *broad income, which* considers changes in taxable income before deductions, which primarily consist of employer administrated pension contributions.

From Panel B, we see the expected downward-sloping curves for both periods, which is consistent with mean reversion being the dominant (but not necessarily the sole) feature of the underlying income process. However, on top of this overall pattern we see a marked difference between the pre-reform and reform periods. In the identification region, where individuals experience a large increase in the marginal net-of tax rate, the growth rate of taxable income in the reform period lies significantly above the growth rate in the pre-reform period (taking into account changes in the overall income growth rate). In contrast, we see essentially the same trend differentials in the two periods in the validation region, where individuals were largely unaffected by the tax reform.

Panel C isolates these changes in income trend differentials and is consistent with the identifying assumption of constant trend differentials, and we observe no significant changes within the validation region. For the individuals in the identification region, in contrast, and we observe changes in trend differentials for taxable income that are strongly increasing in initial income, and assuming that the trend differentials across the entire income distribution would have remained constant in the absence of the reform, Panel C provides causal and non-parametric estimates of the taxable income responses to the 2009-10 tax reform. An assumption that appears reasonable given the constant trend differentials observed in the untreated part of the income distribution.

Finally, in Panel D of Figure 3, we translate observed changes in trend differentials into income elasticities, but instead of estimating a single aggregate elasticity, we explore the heterogeneity across the income distribution by running a 2SLS local linear estimator centered at different points throughout the identification region.

Our implementation of the 2SLS local linear estimator follows a standard local linear regression. For each point of initial income in the identification region ($h \in$ 400,000; 450,000;

...), we separately estimate a weighted 2SLS with the following second stage equation

$$\Delta \ln z_{it} = \gamma_0 + \boldsymbol{\gamma_1}' \boldsymbol{D_{it}^{inc}} + \mu_2 D_{it}^{post} + \gamma_3 \Delta \ln \tau_{it} + \gamma_4 \Delta \ln \tau_{it} \times (z_{it-1} - h) + \nu_{it}, \tag{12}$$

where $\boldsymbol{D_{it}^{inc}}$ is a vector of DKK 10,000 income bin dummies.[20] In this equation we include the marginal net-of-tax rate both by itself and interacted linearly with initial income to capture both the level and slope of the income elasticity centered on h. Both of these terms are endogenous variables ($\Delta \ln \tau_{it}$, $\Delta \ln \tau_{it}(z_{it-1} - h)$) and we instrument these with the corresponding predicted between income changes in the net-of-tax rate ($\triangle \ln \bar{\tau}_{zt-1}^{P}$, $\triangle \ln \bar{\tau}_{zt-1}^{P}(z_{it-1} - h)$).[21]

For weights we first assign individuals within the (less treated) validation region to a control group with a constant weight of 1 across all estimations. In the case of the 2009-10 reform, we select the entire validation region as illustrated by the shaded gray area in Panel D. For the treated individuals in the identification region we assign, separately for each estimation, triangular weights (w) within $\pm$ DKK 50,000 of h computed as

$$w = \max\left(50,000 - |z - h|, 0\right) / 50,000. \tag{13}$$

Hence, with this weighting strategy we keep the control group constant and estimate heterogeneous elasticities by moving the treatment group up through the income distributions, and using this strategy, we estimate elasticities for taxable income that are increasing in initial income from below 0.1 for incomes around 400,000 to more than 0.5 for the top of

[20] The income bin dummies run from DKK 250,000 and are capped at DKK 1,500,000, where the data becomes too thin for separate estimation of dummies. Hence, our regressions on the whole sample include 125 income dummies. Using instead the more fine grid DKK 1,000 dummies (1,250 dummies) does not change the results.

[21] By using $\triangle \ln \bar{\tau}_{zt-1}^{P} = E(\triangle \ln \tau_{it-1}^{p} | z_{it-1})$ as the instrument instead of the actual predicted tax changes, $\triangle \ln \bar{\tau}_{it-1}^{P}$, we only identify the income elasticity from the between income tax variation, which is the tax variation we can validate in our new approach.

the income distribution.[22] [23]

Overall, Figure 3 provides compelling evidence of behavioral responses to the 2009-10 tax reform, however, the substantial effect on taxable income could be driven by a range of factors from "real" labor responses reflecting unobserved efforts, occupational choice, hours worked etc., to avoidance or evasion. To separate out some of these factors, we start by excluding self-employed individuals, who typically have more room to change behavior through, for example, tax planning and retained earnings in the firm (see, e.g., le Maire & Schjerning, 2013). Consistently, we find slightly smaller effects when dropping individuals with income from self-employment, but overall the estimated trend differentials in Panel C and elasticities in Panel D are very similar to those found for the whole population. Next, we consider a broader income measure defined as our taxable income measure (personal income in Table 1) before deductions, which primarily consist of employer administrated pension contributions as discussed in Section 3. Looking at broad income, we find significantly smaller responses across the whole income distribution. Hence, most of the behavioral responses to the 2009-10 reform are likely driven by income shifting from pension contributions to taxable income. The responses in broad income are marginally significant for initial income around DKK 450,000, but insignificant at the top of the distribution.

The 2004 Tax Reform

Next, we turn to the 2004 tax reform, which we analyze following the same steps as above. As mentioned in Section 3, the main feature of the 2004 reform was an increase in the middle tax income threshold, and in Panel A of Figure 4, we see the largest changes in marginal net-of-tax rates for personal income between DKK 290,000 and DKK 360,000.[24]

[22] A finding of heterogeneous elasticities will normally invalidate the elasticity point estimates when a reform affects the entire income distribution with no pure control group but only differences in treatment intensities. Hence, in cases with no pure control groups, heterogeneity analyses such as the one above become particularly important as a way of validating the estimated elasticities. This is less of an issue in our setting, where the individuals in the validation region are only treated by small tax changes.

[23] In Appendix Figure A.I we show how the behavioral responses in taxable income build up over time by looking at 2- and 3-year differences in addition to the 4-year differences considered here.

[24] Similar to the abolitiont of the middle tax in the 2009-10 reform, the increased threshold also created within income tax variation due to the joint taxation of couples.

We define this interval as the identification region with validation regions both above and below.

In Panels B and C of Figure 4 we examine the income trend differentials across the income distribution for the pre-reform period (1999-2003) and the reform period (2003-07). Panel B reveals the expected downward sloping pattern consistent with mean reversion, but notably we observe essentially no changes in income trend across the identification region from the pre-reform to the reform period. The only changes in trend differentials that stand out from Panel B are for initial income between DKK 200,000 and DKK 225,000, which is difficult to reconcile with being responses to the changes in taxation. The same is visible in Panel C where, in addition, to taxable income we also consider the changes excluding self-employed and changes in broad income. For broad income, we find slightly higher income trends for the top of the income distribution (above DKK 500.000), but similar to the change in income trends in the bottom, these are difficult to reconcile with the changes in taxation.

In Panel D, we translate the estimated trend differentials in Panel C into elasticities using the local linear 2SLS estimator described in equations (12)-(13) above. However, as the reform creates only a narrow identification region, it is unfeasible to explore heterogeneity within this region so we only estimate a single point elasticity. A more relevant question when estimating elasticities from the 2004 reform is who to include in the control group. Based on the discussion above, and contrary to our analysis of the 2009-10 reform, we do not include the entire validation region. Instead, we select the grey income intervals immediately adjacent to the identification regions in Panel D as control groups. In this way we avoid letting the changes in trend differentials observed at the very top and bottom of the income distribution affect the estimates, and from this we obtain a precisely estimated zero for both taxable and broad income. This result is not necessarily surprising given that the 2004 reform was smaller, and hence, less likely to prompt individuals to respond (Chetty, 2012).

Results from the Standard Estimation Approach

Above, we analyzed the 2004 and 2009-10 reforms using our new approach, which, in particular, allowed us to inspect the changes in trend differentials graphically and, hence, to validate the identifying assumption of constant trend differentials in the absence of tax reforms. To highlight the importance of such validation exercises, we examine the two reforms using the standard approach. More specifically, we run the following 2SLS regression separately for the 2004 and 2009-10 tax reforms

$$\Delta \ln z_{it} = \gamma_0 + \boldsymbol{\gamma_1}' \boldsymbol{D_{it}^{inc}} + \mu_2 D_{it}^{post} + \gamma_3 \Delta \ln \tau_{it} + \nu_{it}, \tag{14}$$

where $\boldsymbol{D_{it}^{inc}}$ is a vector of DKK 10,000 income bin dummies. This equation contains one endogenous variable ($\Delta \ln \tau_{it}$), which we instrument with the corresponding predicted between income changes in the marginal net-of-tax rate ($\triangle \ln \bar{\tau}_{zt-1}^{P}$). We run the regressions on the same data and consider the same four year periods and outcomes as in Figure 3 and 4 above, and we present the results in Table 4.

Considering first the results for the 2009-10 reform in columns (4)-(6), we estimate an average elasticity of just above 0.2 for taxable income with (column 4) or without (column 5) self-employed individuals, while the elasticity for broad income is close to zero but still positive and significant. These estimates are well in line with the results from our graphical analysis above, which is not surprising given that we found that trend differentials were constant across the entire validation region in the case of the 2009-10 reform.

Next, considering the 2004 reform in columns (1)-(3), we find elasticities that very similar to results from the 2009-10 reform. However, in our graphical analysis in Figure 4 we saw no changes in trend differentials in connection with changes in tax treatment. Instead, the results we obtain from the standard approach are driven by the changes in trend differentials observed at the very top and bottom of the income distribution, which is less likely to reflect behavioral responses to the tax reform. Hence, in this case, the estimates from the standard approach are most likely biased.

In Appendix E, we analyze two additional tax reforms and essentially arrive at the same conclusion as for the 2004 and 2009-10 reforms. For a reform in 1987, we find support

for the identifying assumption of constant trend differentials and positive responses for taxable income, while for the reform in 1994, we find clear violations of the identifying assumption.

5 Conclusion

Behavioral responses to taxes are key inputs in the design of economic policy and serve as evidence of behavioral parameters in economic models more broadly. In this paper, we revisited the identification of behavioral responses to tax reforms and developed a new approach that allows for graphical validation of key identifying assumptions and representation of treatment effects.

Considering stereotypical tax reforms, which change tax rates in one part of the income distribution, while keeping them constant in the rest of the distribution, we show that the state-of-the-art estimation strategy employed by, for example, Gruber & Saez (2002) and Kleven & Schultz (2014), in essence, relies on an assumption that trend differences in income across the income distribution remain constant in the absence of reforms. Similar to the pre-trend validation of differences-in-differences studies, this identifying assumption of constant trend differentials can be validated by comparing the evolution of income in untreated parts of the income distribution over time.

We illustrate the importance of our new validation approach by studying a number of tax reforms in Denmark, with our main applications being the 2004 and 2009-10 reforms. Analyzing both reforms through the lens of the standard estimation approach we find very similar average income elasticities in the order 0.2 for taxable income and 0.01 for broad income. However, comparing the trend differentials for the pre-reform and reform periods, it is clear that only the results from the 2009-10 reform are likely to be causal. For this reform, we find that the income trend differentials remained stable in the untreated bottom part of the income distribution, which is consistent with the identifying assumption, while for the treated upper part of the income distribution, we find significantly higher income growth that is strongly increasing in initial income and somewhat larger in the medium run than in the short run.

In contrast, we find no changes in trend differentials around the changes in tax treatment created by the 2004 reform. Instead, the elasticity estimates for this reform are driven by changes in trend differentials well within the control group, which are most likely unrelated to the reform. Once we account for these changes, we obtain a precisely estimated zero for both taxable and broad income.

References

ANDERSEN, HENRIK YDE. 2018. Do Tax Incentives for Saving in Pension Accounts Cause Debt Accumulation? Evidence from Danish Register Data. *European Economic Review*, **106**.

ARELLANO, MANUEL, & BOND, STEPHEN. 1991. Some Tests of Specification for Panel Data Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, **58**(2).

AUTEN, GERALD, & CARROLL, ROBERT. 1999. The Effect of Income Taxes on Household Income. *Review of Economics and Statistics*, **81**(4).

CESARINI, DAVID, LINDQVIST, ERIK, NOTOWIDIGDO, MATTHEW J., & ÖSTLING, ROBERT. 2017. The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries. *American Economic Review*, **107**(12).

CHETTY, RAJ. 2012. Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply. *Econometrica*, **80**(3).

CHETTY, RAJ, FRIEDMAN, JOHN N., OLSEN, TORE, & PISTAFERRI, LUIGI. 2011a. Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics*, **126**(2).

CHETTY, RAJ, GUREN, ADAM, MANOLI, DAY, & WEBER, ANDREA. 2011b. Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins. *In: American Economic Review: Papers & Proceedings*, vol. 101.

CHETTY, RAJ, FRIEDMAN, JOHN N., LETH-PETERSEN, SØREN, NIELSEN, TORBEN HEIEN, & OLSEN, TORE. 2014. Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark. *The Quarterly Journal of Economics*, **129**(3).

CLEMENS, JEFFREY, & WITHER, MICHAEL. 2019. The Minimum Wage and the Great Recession: Evidence of Effects on the Employment and Income Trajectories of Low-Skilled Workers. *Journal of Public Economics*, **170**.

CURRIE, JANET, & FALLICK, BRUCE C. 1996. The Minimum Wage and the Employment of Youth Evidence from the NLSY. *The Journal of Human Resources*, **31**(2).

DEVEREUX, MICHAEL P., LIU, LI, & LORETZ, SIMON. 2014. The Elasticity of Corporate Taxable Income: New Evidence from UK Tax Records. *American Economic Journal: Economic Policy*, **6**(2).

DUSTMANN, CHRISTIAN, LINDNER, ATTILA, SCHÖNBERG, UTA, UMKEHRER, MATTHIAS, & VOM BERGE, PHILIPP. 2019. Reallocation Effects of the Minimum Wage: Evidence From Germany. *Working Paper*.

EISSA, NADA, & LIEBMAN, JEFFREY B. 1996. Labor Supply Response to the Earned Income Tax Credit. *The Quarterly Journal of Economics*, **111**(2).

FACK, GABRIELLE, & LANDAIS, CAMILLE. 2016. The Effect of Tax Enforcement on Tax Elasticities: Evidence from Charitable Contributions in France. *Journal of Public Economics*, **133**.

FELDSTEIN, MARTIN. 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).

FELDSTEIN, MARTIN. 1999. Tax Avoidance and the Deadweight Loss of the Income Tax. *Review of Economics and Statistics*, **81**(4).

GARICANO, LUIS, LELARGE, CLAIRE, & VAN REENEN, JOHN. 2016. Firm Size Distortions and the Productivity Distribution: Evidence from France. *American Economic Review*, **106**(11).

GIERTZ, SETH H. 2015. The Elasticity of Taxable Income over the 1980s and 1990s. *National Tax Journal*, **60**(4).

GOOLSBEE, AUSTAN. 2000. What Happens When You Tax the Rich? Evidence from Executive Compensation. *Journal of Political Economy*, **108**(2).

GORDON, ROGER, & SLEMROD, JOEL. 1998. Are "Real" Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases? *NBER Working Paper No. 6576*.

GRUBER, JONATHAN, & SAEZ, EMMANUEL. 2002. The Elasticity of Taxable Income: Evidence and Implications. *Journal of Public Economics*, **84**(1).

IMBENS, GUIDO W, RUBIN, DONALD B, & SACERDOTE, BRUCE I. 2001. Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players. *American Economic Review*, **91**(4).

JAKOBSEN, KATRINE, JAKOBSEN, KRISTIAN, KLEVEN, HENRIK, & ZUCMAN, GABRIEL. 2020. Wealth Taxation and Wealth Accumulation: Theory and Evidence from Denmark. *Quarterly Journal of Economics*, **135**(3).

KLEVEN, HENRIK JACOBSEN. 2020. The EITC and the Extensive Margin: A Reappraisal. *NBER Working Paper No. 26405*.

KLEVEN, HENRIK JACOBSEN, & SCHULTZ, ESBEN ANTON. 2014. Estimating Taxable Income Responses Using Danish Tax Reforms. *American Economic Journal: Economic Policy*, **6**(4).

KLEVEN, HENRIK JACOBSEN, KNUDSEN, MARTIN B., KREINER, CLAUS THUSTRUP, PEDERSEN, SØREN, & SAEZ, EMMANUEL. 2011. Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, **79**(3).

KOPCZUK, WOJCIECH. 2005. Tax Bases, Tax Rates and The Elasticity of Reported Income. *Journal of Public Economics*, **89**(11-12).

KREINER, CLAUS THUSTRUP, LETH-PETERSEN, SØREN, & SKOV, PEER EBBESEN. 2016. Tax Reforms and Intertemporal Shifting of Wage Income: Evidence from Danish Monthly Payroll Records. *American Economic Journal: Economic Policy*, **8**(3).

KREINER, CLAUS THUSTRUP, LETH-PETERSEN, SØREN, & SKOV, PEER EBBESEN. 2017. Pension Saving Responses to Anticipated Tax Changes: Evidence from Monthly Pension Contribution Records. *Economics Letters*, **150**.

KUMAR, ANIL, & LIANG, CHE-YUAN. 2020. Estimating Taxable Income Responses with Elasticity Heterogeneity. *Journal of Public Economics*, **188**.

LE MAIRE, DANIEL, & SCHJERNING, BERTEL. 2013. Tax Bunching, Income Shifting and Self-employment. *Journal of Public Economics*, **107**.

MEYER, BRUCE D., & ROSENBAUM, DAN T. 2001. Welfare, The Earned Income Tax Credit, and The Labor Supply of Single Mothers. *Quarterly Journal of Economics*, **116**(3).

MIAN, ATIF, & SUFI, AMIR. 2010. Household Leverage and the Recession of 2007-09. *IMF Economic Review*, **58**(1).

PIRTTILÄ, JUKKA, & SELIN, HÅKAN. 2011. Income Shifting within a Dual Income Tax System: Evidence from the Finnish Tax Reform of 1993. *Scandinavian Journal of Economics*, **113**(1).

SAEZ, EMMANUEL. 2001. Using Elasticities to Derive Optimal Income Tax Rates. *Review of Economic Studies*, **68**(1).

SAEZ, EMMANUEL. 2003. The Effect of Marginal Tax Rates on Income: A Panel Study of 'Bracket Creep'. *Journal of Public Economics*, **87**(5-6).

SAEZ, EMMANUEL, SLEMROD, JOEL, & GIERTZ, SETH H. 2012. The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature*, **50**(1).

SEIM, DAVID. 2017. Behavioral Responses to Wealth Taxes: Evidence from Sweden. *American Economic Journal: Economic Policy*, **9**(4).

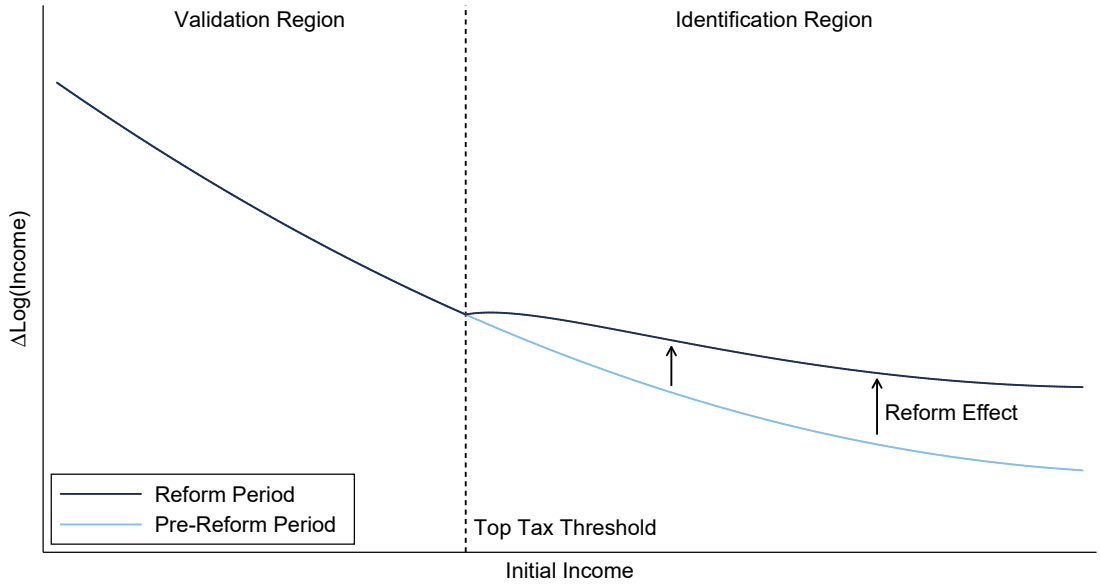
SLEMROD, JOEL. 1995. Income Creation or Income Shifting? Behavioral Responses to the Tax Reform Act of 1986. *The American Economic Review*, **85**(2).

SLEMROD, JOEL, & KOPCZUK, WOJCIECH. 2002. The Optimal Elasticity of Taxable Income. *Journal of Public Economics*, **84**(1).

WEBER, CAROLINE E. 2014. Toward Obtaining a Consistent Estimate of the Elasticity of Taxable Income Using Difference-in-Differences. *Journal of Public Economics*, **117**.

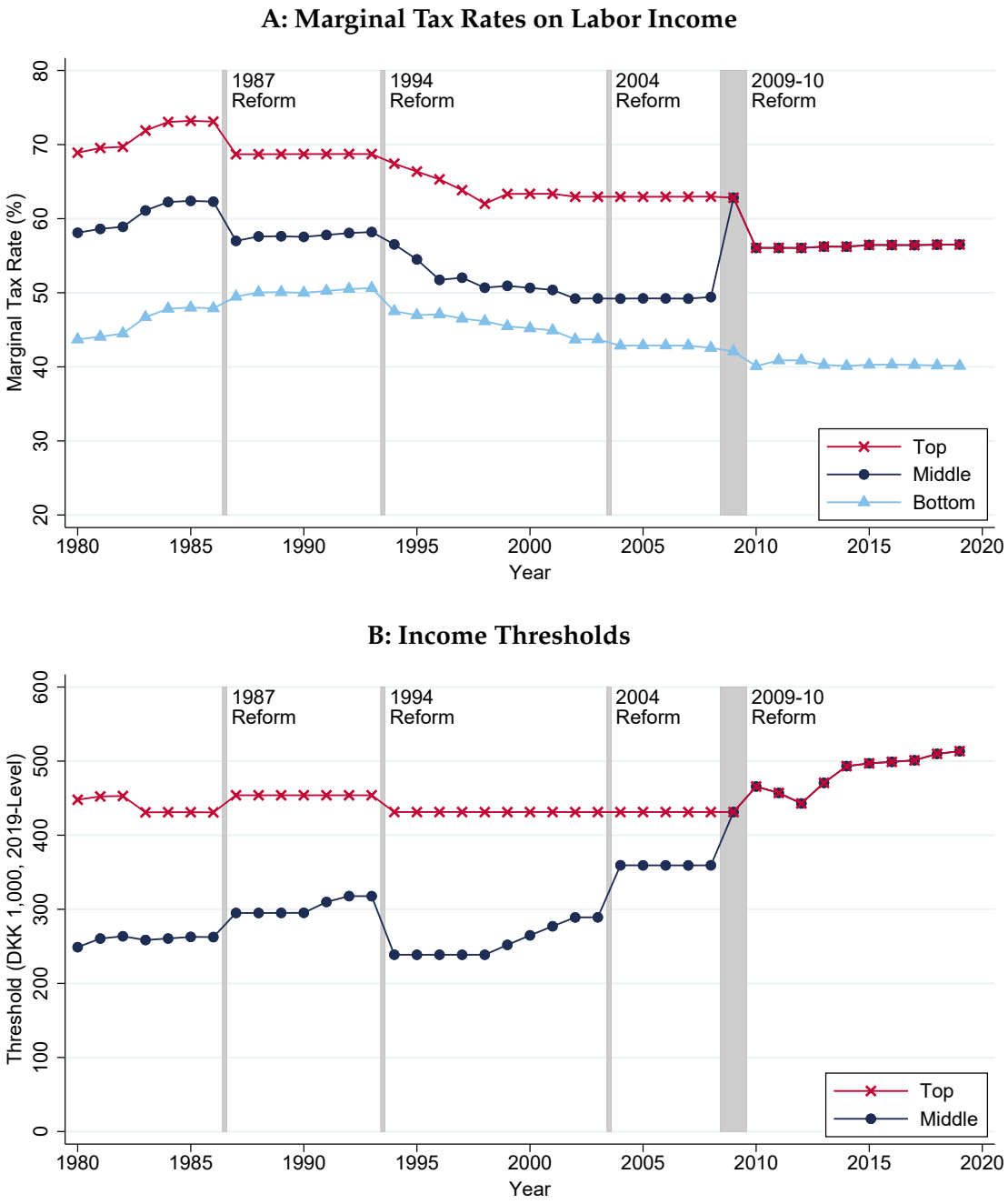
Figures and Tables

Figure 1: Illustration of the Identification and Validation Strategy



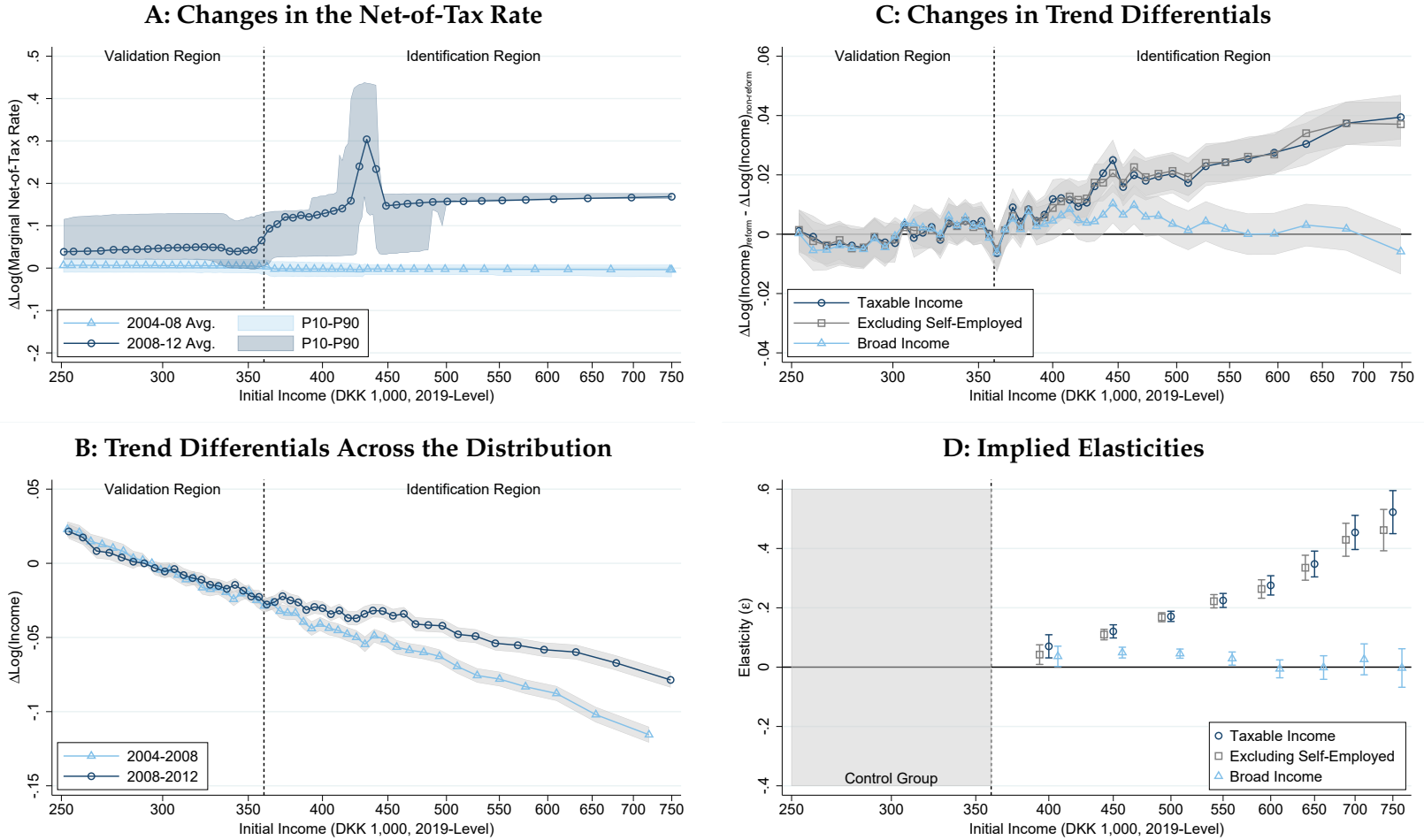
Notes: The figure shows $E\left(\triangle \ln z_{it} | z_{it-1}\right)$, i.e., the changes in log income across the income distribution for two time periods: a pre-reform period, where the tax system remains stable and a reform period, where the top tax is reduced. The negative relationship between $E\left(\triangle \ln n_{it} | z_{it-1}\right)$ and z_{it-1} is consistent with mean reversion being the dominate - but not necessarily the only - feature of the underlying income process. Under the assumption that this pattern would have remained constant (relative to the overall growth in the economy) in absence of the reform, we can identify the reform effect from the differences between the reform and pre-reform periods for the population with initial income above the top tax threshold (the identification region). In contrast, we should observe no differences for the part of the population with initial income below the top tax threshold (the validation region). Thus, comparing the reform and pre-reform periods in this part of the income distribution acts as a placebo test for the validity of the identifying assumption.

Figure 2: Key Features of the Danish Tax Schedule 1980-2020



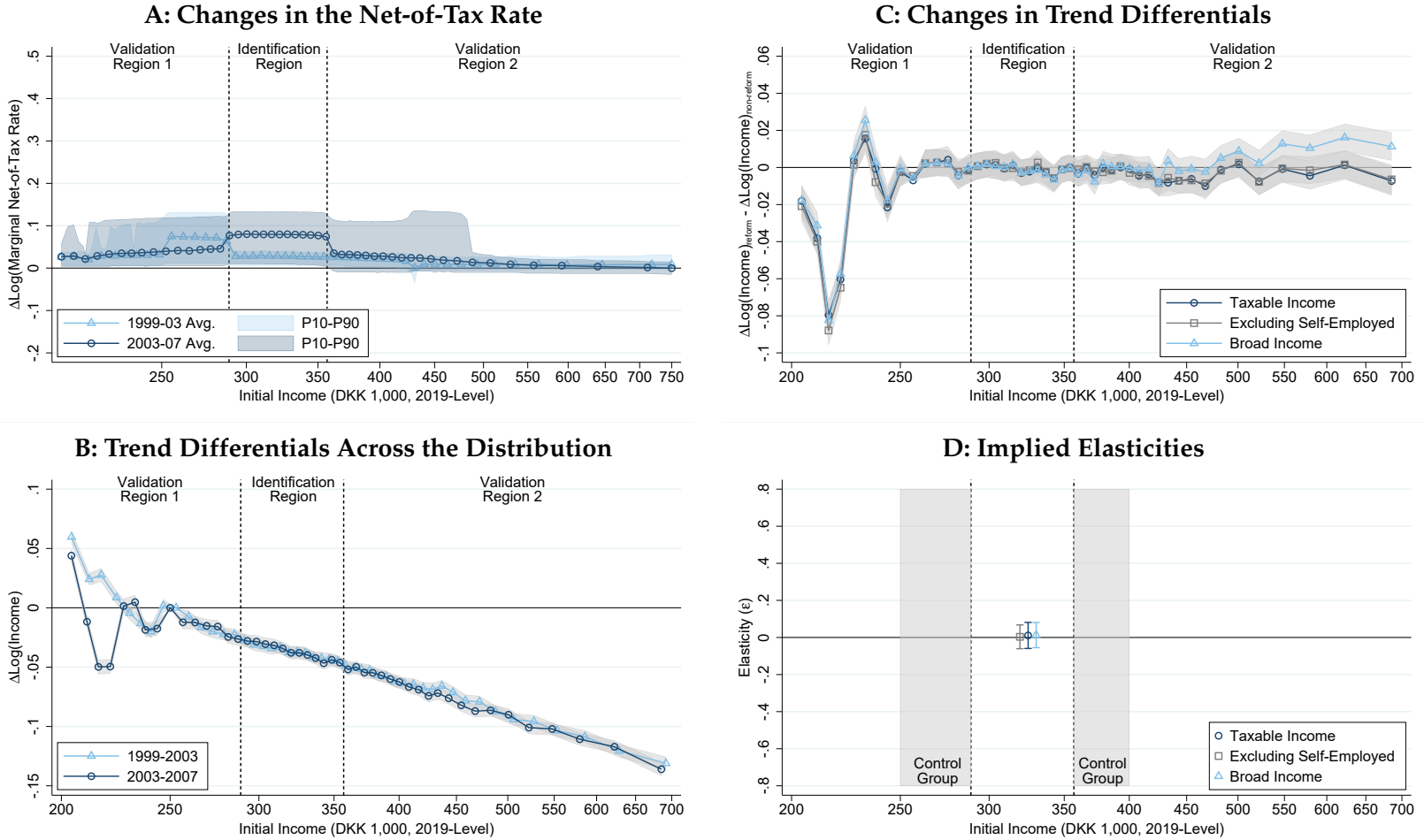
Notes: The figure shows the key features of the Danish tax system 1980-2020. Panel A shows the marginal tax rates of the three (main) income tax brackets, while Panel B shows the income thresholds. The thresholds are deflated to 2019-levels using the implied wage indexes in the tax code. Hence, changes in the thresholds reflect active policy decisions. The figure highlights the four major reforms that we analyze. USD 1 $\approx$ DKK 6.8.

Figure 3: Identifying Income Responses Following the 2009-10 Tax Reform



Notes: Panel A shows the predicted changes in the log net-of-tax rate (τ^{p}_{it-1}) from 2004 to 2008 (the pre-reform period) and 2008 to 2012 (the reform period). The curves show the average changes within each income bin and the shaded areas show the 10th and 90th percentile ranges. Panel B shows the estimated income trend differentials using equation (10) for 2004-08 and 2008-12 relative to the average growth rate for incomes around DKK 300,000. Panel C shows the estimated changes in trend differentials based on equation (11) for different samples and income concepts. The use of different income concepts only affects the dependent variable (y-axis). Initial income refers to personal income as defined in Table 1 and is measured in 2004 and 2008, and is thus pre-reform. Panel D shows the implied elasticities over in the identification region using the 2SLS local linear estimation described in equation (12). The grey shaded area illustrates the interval included as the control group in the elasticity estimations. USD 1 $\approx$ DKK 6.8.

Figure 4: Identifying Income Responses Using the 2004 Tax Reform



Notes: Panel A shows the predicted changes in the log net-of-tax rate (τ^p_{it-1}) from 1999 to 2003 (the pre-reform period) and 2003 to 2007 (the reform period). The curves show the average changes within each income bin and the shaded areas show the 10th and 90th percentile ranges. Panel B shows the estimated income trend differentials using equation (10) for 1999-2003 and 2003-07 relative to the average growth rate for incomes around DKK 250,000. Panel C shows the estimated changes in trend differentials based on equation (11) for different samples and income concepts. The use of different income concepts only affects the dependent variable (y-axis). Initial income refers to personal income as defined in Table 1 and is measured in 1999 and 2003, and is thus pre-reform. Panel D shows the implied elasticities over in the identification region using the 2SLS local linear estimation described in equation (12). The grey shaded area illustrates the interval included as the control group in the elasticity estimations. USD 1 $\approx$ DKK 6.8.

Table 1: Income Concepts in the Danish Income Tax System

Income concept	**Acronym**	**Main items included**
Labor Market Income	LI	Salary, wages, honoraria, fees, bonuses, fringe benefits, business earnings
Personal Income	PI	LI + transfers, certain pension payouts, received alimony − labor market contribution, certain pension contributions[a)]
Capital Income	CI	Interest income, rental income, business capital income − interest on debt (mortgages, bank loans, credit cards, student loans)
Stock Income	SI	Dividends and realized capital gains from shares
Itemized Deductions	ID	Commuting, union fees, unemployment insurance contributions, other work expenditures, charitable donations, paid alimony
Taxable Income	TI	PI + CI − ID

a) Employer administrated contributions to annuity pension were completely tax deductible until 2010. After 2010 a cap of DKK 100,000 on contributions to expiring annuities was introduced, as described in the text.

Table 2: Tax Bases and Tax Rates over Time

		2004 Reform		2009-10 Reform	
Tax type	Base[e)]	Pre-Reform Rate (%)	Post-Reform Rate (%)	Pre-Reform Rate (%)	Post-Reform Rate (%)
Labor Market Contribution[a)]	LI	8.0	8.0	8.0	8.0
EITC[b)]	LI	0.0	2.5	4.0	4.3
Regional Tax[c)]	TI	32.6	32.6	33.5	32.9
Bottom Tax	PI + [CI>0]	5.5	5.5	5.5	3.7
Middle Tax	PI + [CI>0]	6.0	6.0	6.0	0.0
Top Tax	PI + [CI>0]	15.0	15.0	15.0	15.0
Marginal tax ceiling[d)]	-	59.0	59.0	59.0	51.5

a) Because labor market income enters the other tax bases net of the labor market contribution, the effective tax rate on labor income equals the statutory tax rate times (1 – the labor market contribution rate).
b) The Danish EITC is treated as an itemized deduction in taxable income (TI). Hence, the tax value of the earned income tax credit (EITC) is the EITC rate times the marginal tax on taxable income.
c) The regional tax includes municipal taxes and health contributions. The regional tax rate in the table is the average across all municipalities.
d) If the sum of all regional and national tax rates (excluding the stock income tax and the labor market contribution) exceeds the specified ceiling, the top tax is adjusted downward until the marginal tax rate equals the ceiling.
e) There have been a few minor changes to the tax bases over time related to the treatment of capital income.

Table 3: Sample and Initial Descriptive Statistics

	Pre-Reform				Reform			
Panel A: 2004 Reform	**P25**	**P50**	**Mean**	**P75**	**P25**	**P50**	**Mean**	**P75**
Broad Income (1,000 DKK)	296	375	419	470	301	384	430	484
Personal Income (1,000 DKK)	279	350	389	437	276	345	386	433
Labor Income (1,000 DKK)	270	377	387	482	261	370	377	476
Capital Income (1,000 DKK)	-42	-16	-21	-0	-49	-21	-27	-2
Age	32	41	42	51	34	43	43	53
Self-Employed (%)	-		14	-	-		13	-
Married (%)	-		66	-	-		68	-
Number of Observations		2,240,332				2,295,922		
Panel B: 2009-10 Reform	**P25**	**P50**	**Mean**	**P75**	**P25**	**P50**	**Mean**	**P75**
Broad Income (1,000 DKK)	346	417	473	517	359	436	498	546
Personal Income (1,000 DKK)	313	373	419	457	322	386	434	473
Labor Income (1,000 DKK)	318	398	423	497	326	409	437	513
Capital Income (1,000 DKK)	-48	-24	-27	-3	-63	-31	-39	-4
Age	35	44	44	54	36	45	45	55
Self-Employed (%)	-		15	-	-		13	-
Married (%)	-		72	-	-		72	-
Number of Observations		2,097,862				2,240,833		

Notes: Monetary units are deflated to 2019-levels using the implied wage indexes in the tax code. (USD 1 $\approx$ DKK 6.8). Personal income, labor income, and capital income are defined in Table 1. Broad income corresponds to personal income before deductions.

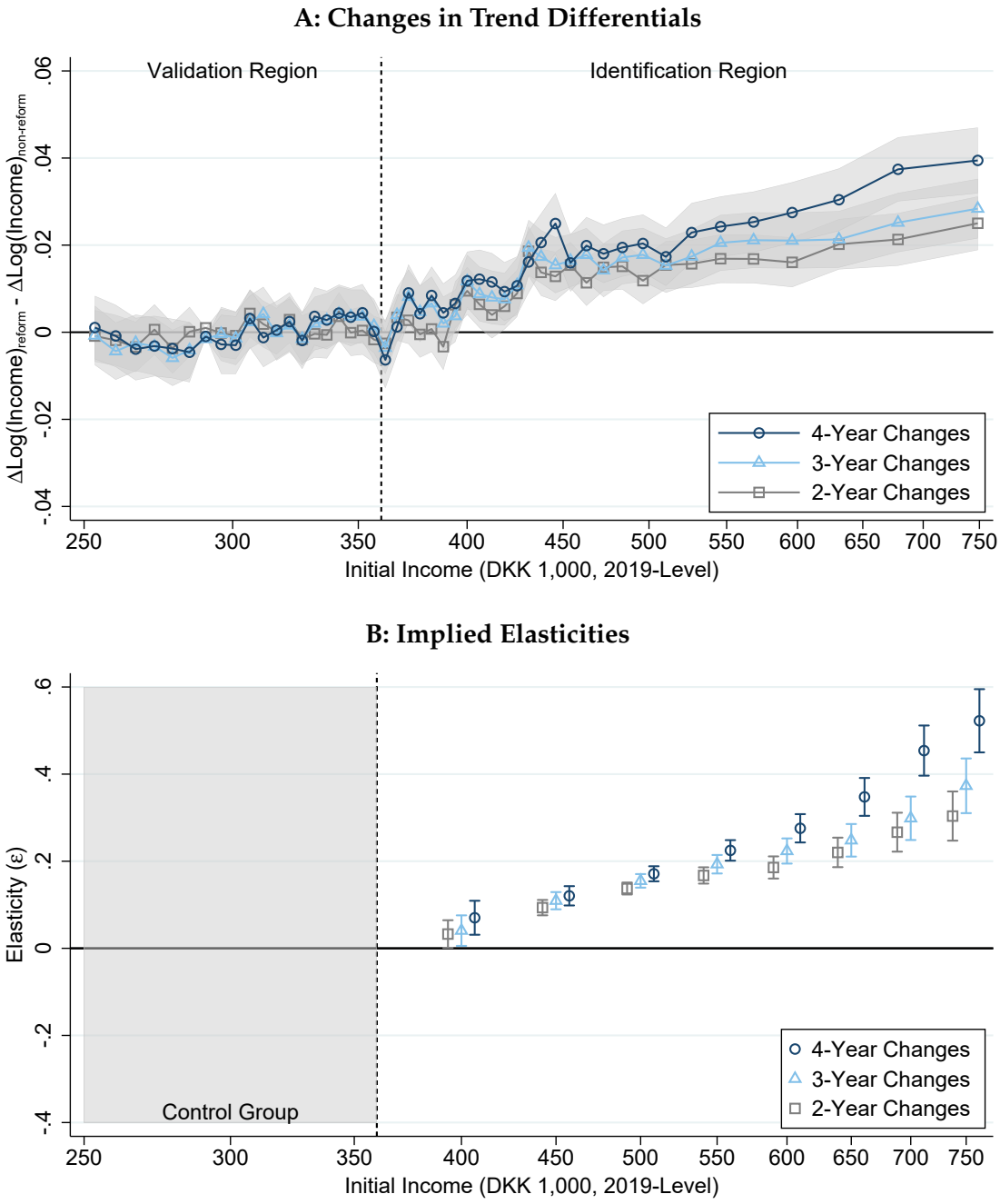
Table 4: Standard Approach Income Elasticity Estimates from the 2004 Reform and the 2009-10 Reforms

	2004 Reform			**2009-10 Reform**		
	Taxable Income	Excluding Self-Employed	Broad Income	Taxable Income	Excluding Self-Employed	Broad Income
	(1)	(2)	(3)	(4)	(5)	(6)
Second Stage						
$\triangle \log(\tau_{it})$	0.218	0.288	0.062	0.227	0.202	0.016
	(0.046)	(0.041)	(0.039)	(0.009)	(0.007)	(0.007)
First Stage						
$\triangle \log(\tau_{it}^{P})$	0.314	0.342	0.342	0.661	0.726	0.726
	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)
Income Controls	DKK 10,000 Bins					
Time Periods	1999-2003	1999-2003	1999-2003	2004-2008	2004-2008	2004-2008
	2004-2008	2004-2008	2004-2008	2008-2012	2008-2012	2008-2012
Obs.	4,536,254	3,906,266	3,906,266	4,338,695	3,732,168	3,732,168

Notes: The table summarizes the results of the standard estimation approach described in equation (14) for the 2004 and 2009-10 reforms. Robust standard errors in parentheses.

A Additional Figures and Tables

Figure A.I: Dynamic Income Responses to the 2009-10 Tax Reform



Notes: This figure shows the dynamics of the responses to the 2009-10 reform. Panel A shows the estimated changes in trend differentials based on equation (11) for 2-year, 3-year, and 4-year changes in taxable income between the pre-reform period and the reform period. The baseline years are always 2004/2008. Hence, the 2-year differences compare the change in income for 2008-10 to the change in income for 2004-06, and the 3-year differences are changes from 2008-11 to 2004-07. Panel B shows the implied elasticities over the income distribution estimated using the 2SLS local linear estimation described in equation (12).

B Reinterpreting the Weber (2014) Approach

In this section, we consider the more recent work by Weber (2014) and show how we can reinterpret this work in light of our new approach laid out in Section 2 in the main paper. To reduce the problem of mean reversion, Weber (2014) suggests using further lags of past taxable income (z_{it-1-k}) when computing the predicted changes in marginal net-of-tax rates. This is likely to work, as any transitory component in potential income (by definition) will play out over time and hence, not affect changes in potential income sufficiently far into the future. Mathematically, we should, therefore, expect the exclusion restriction to hold when past taxable income is measured sufficiently far back:[25]

$$\text{cov}\left(\triangle \ln \tau^{p}_{it-1-k}, \triangle \ln n_{it}\right) \to 0 \quad \text{for} \quad k \to \infty. \tag{15}$$

Related to our new approach, there are two things worth highlighting in the approach presented by Weber (2014). First, if the exclusion restriction (15) does hold, the effect of a tax reform on taxable income can be identified from the changes in taxable income from the reform period alone. There is no need for a pre-reform period to map out the underlying income process, as this is orthogonal to the variation the simulated instruments when (15) holds.

Second, the validity of the exclusion restriction can, however, be tested in a way very similar to the one presented in the main paper. Because the predicted tax rates here are a function of lagged taxable income z_{it-1-k}, we can validate the exclusion restriction (15) by examining the relationship between z_{it-1-k} and $E\left(\triangle \ln z_{it} | z_{it-1-k}\right)$ in parts of the income distribution untreated by tax reforms.

Graphically, the validation can be shown in figures similar to Figure 1 in the main paper, but with z_{it-1-k} on the x-axis. If the exclusion restriction (15) is fulfilled, we should expect a flat curve for the untreated parts of the income distribution, as illustrated in Figure A.II. Put differently, if mean reversion is the only source of trend differentials across the

[25] In principle, it is not a sufficient condition that the covariance in equation (15) is converging to zero for the bias of the 2SLS estimate to converge to zero. For that, the covariance needs to converge to zero *faster* than the first stage cov $\left(\triangle \ln \tau^{p}_{it-1-k}, \triangle \ln \tau_{it}\right)$.

income distribution and we assume that all transitory shocks to income no longer affect the income process after year k, then we should expect no differences in income trends across untreated subsamples of the population defined by their income k years ago.

While the above the validation can be perfomed using only the reform period, we can provide further validation by using one or more pre-reform periods. Given the exclusion restriction indeed holds, we would expect a flat relationship between $E\left(\triangle \ln z_{it} | z_{it-1-k}\right)$ and z_{it-1-k} for the entire income distribution in pre-reform periods, as illustrated in Figure A.II. On the other hand, if this is not the case, we are essentially back to our new approach, where we can identify the taxable income responses by comparing changes in income trend differentials across treated and untreated parts of the income distribution with the only difference being that we define the income distribution using further lags of taxable income.

We illustrate how the work by Weber (2014) relates to our new approach by applying the method described above to the tax reform in 2009-10. In Figure A.III and Table A.I, we investigate the income trend differentials across the income distribution in the pre-reform (2004-08) and reform (2008-12) periods, using lags of past taxable income (z_{it-1-k}) to group individuals. Formally, we run a regression similar to equation (10) separately for each period t, where $\boldsymbol{D_{it}^{inc}}$ is a vector of initial income bin dummies measured in $t-k$ for $k \in \{0,1,2,3\}$.

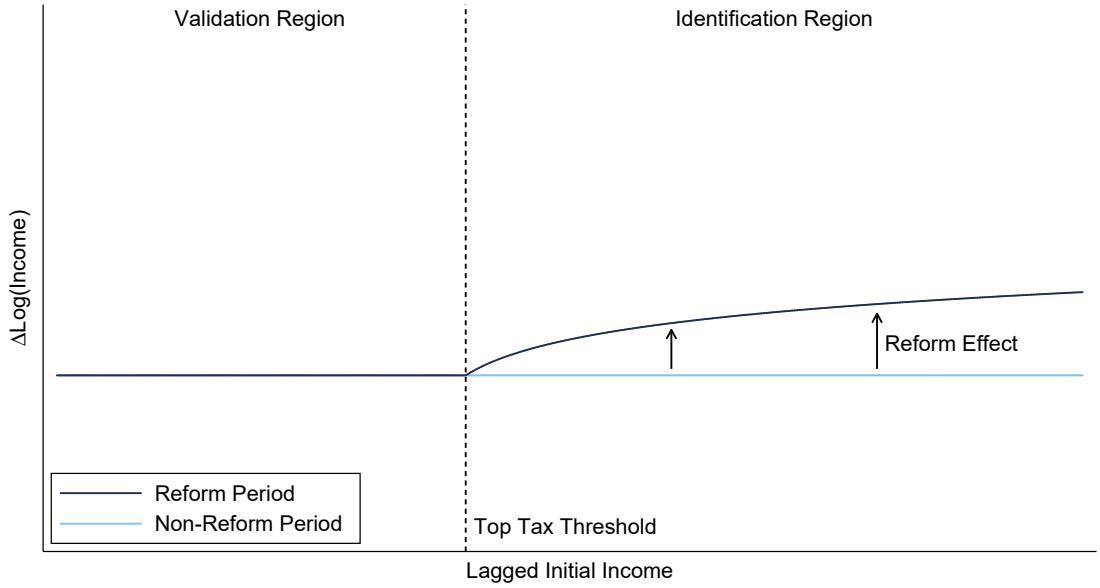
Plotting $\boldsymbol{\beta_1^t}$ for the two periods produces the empirical equivalent of our stylized Figure A.II.[26] Consistent with Weber (2014), the mean reversion problem becomes less pronounced when using further lags of initial income as the overall slope of both curves becomes flatter and the 2SLS estimates without income controls become less negative. However, even with a lag length of three years, we still find differences in income trends across the income distribution.

Looking at the regression including income controls in Table A.I, where we use the pre-reform period to control for the (remaining) income trend differentials, we see positive

[26] For each lag length (k), we select the sample as individuals who have taxable income above DKK 250,000 in both $t-1$ and $t-1-k$. In principle, selecting individuals based on their income in multiple years could in itself reduce the problem of mean reversion. In our case, the changing sample only has minor effects on the estimates.

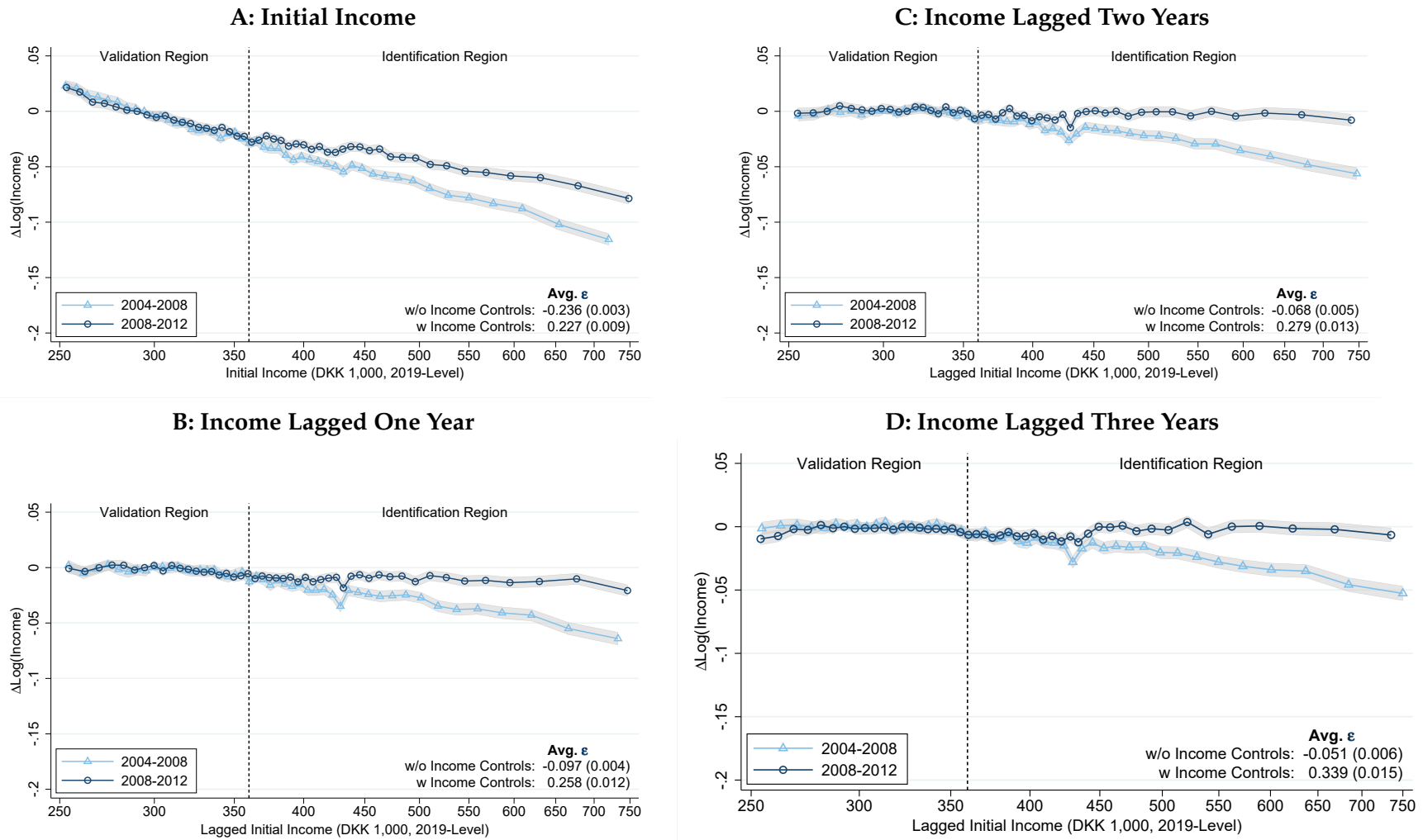
elasticity estimates of the same order as our baseline approach but increasing in the lag length. Mechanically, this pattern emerges due to the first stage estimates falling at a faster rate than the reduced form estimates. More conceptually, we can think of this pattern as being driven by a changing set of compliers. The compliers in these estimations are the set of individuals who stay within their income bracket and, hence, are treated as prescribed by the tax reform and using further lags we are selecting a decreasing set of individuals who stay within their tax bracket for longer.

Figure A.II: Identification and Validation of the Weber (2014) Estimation Strategy



Notes: This figure shows $E\left(\triangle \ln z_{it} | z_{it-1-k}\right)$, i.e., the change in taxable income across the lagged income distribution for two time periods: a pre-reform period, where the tax system remains stable and a reform period, where the top tax is reduced. The flat relationship between $E\left(\triangle \ln n_{it} | z_{it-1-k}\right)$ and z_{it-1-k} in the validation period is consistent with the identifying assumption in Weber (2014). That is, the use of further lags of income to construct the simulated instruments fully resolves the problem of mean reversion. Under this assumption, the reform effect can simply be estimated as the difference in income growth between treated and untreated parts of the income distribution. However, with a pre-reform period available, the identifying assumption can be validated even further, as illustrated in the figure.

Figure A.III: Identifying Income Responses Using Lagged Initial Income



Notes: The figure replicates Figure 3, Panel B for the 2009-10 reform by showing $E\left(\triangle \ln z_{it} | z_{it-1-k}\right)$ for $k \in \{0,1,2,3\}$. It plots the estimated income trend differentials using equation (10) for 2004-08 and 2008-12 relative to the growth rate for income around DKK 300,000. Initial income is measured in $2004-k$ and $2008-k$ and income always refers to personal income as defined in Table 1. USD 1 $\approx$ DKK 6.8.

Table A.I: Summary Income Elasticities Using Lagged Initial Income

	Initial Income	**Lagged Initial Income**		
	$k = 0$	$k = 1$	$k = 2$	$k = 3$
	(1)	(2)	(3)	(4)
Panel A: Regressions Without Income Controls				
Second Stage				
$\triangle \log(\tau_{it})$	-0.236	-0.097	-0.068	-0.051
	(0.003)	(0.004)	(0.005)	(0.006)
First Stage				
$\triangle \log(\tau_{it}^{P})$	1.375	0.951	0.827	0.748
	(0.001)	(0.002)	(0.002)	(0.002)
Income Controls		None		
Time Periods		2008-2012		
Obs.	2,240,833	2,032,318	1,932,857	1,843,289
Panel B: Regressions with Income Controls				
Second Stage				
$\triangle \log(\tau_{it})$	0.227	0.258	0.279	0.339
	(0.009)	(0.012)	(0.013)	(0.015)
First Stage				
$\triangle \log(\tau_{it}^{P})$	0.661	0.505	0.461	0.413
	(0.002)	(0.002)	(0.002)	(0.002)
Income Controls		DKK 10,000 Bins		
Time Periods		2004-2008 & 2008-2012		
Obs.	4,338,695	3,780,870	3,774,774	3,628,103

Notes: The table summarizes the estimated elasticity of taxable income from the 2009-10 reform using the 2SLS estimator similar to equation (14) in the main paper, but using lags of past taxable income (z_{it-1-k}) when computing the predicted changes in marginal net-of-tax rates. Panel A shows the estimated elasticities using only the reform-period and no income controls. Panel B shows the estimated elasticities using both the pre-reform and reform periods and including controls for (lag) initial income. For each lag length (k), we select the sample as individuals who have taxable income above DKK 250,000 in both $t-1$ and $t-1-k$. The changing sample only has minor effects on the estimates.

C Within Income Tax Variation

In this section we consider the estimation of behavioral responses using *within* income tax variation $\left(\triangle \ln \tau^p_{it-1} - \triangle \ln \bar{\tau}^p_{zt-1}\right)$ as defined in equation (7). Restating the definition of the predicted changes in marginal net-of-tax rates

$$\triangle \ln \tau^p_{it-1} = \ln\left(1 - T'_t(z_{it-1}; x_{it-1})\right) - \ln\left(1 - T'_{t-1}(z_{it-1}; x_{it-1})\right), \tag{16}$$

we can think of within income variation as being driven by individual differences in x_{it-1} in contrast to the between income variation driven by differences in z_{it-1}.

Using within income variation we have, by definition, variation in marginal net-of-tax rates conditional on initial income; hence, we can estimate behavioral responses using the standard Gruber & Saez (2002) approach on a single reform period.[27] However, the fact that we execute the estimation does not guarantee consistency of the estimates. For that, we still need the exclusion restriction

$$\text{cov}\left(\triangle \ln \tau^p_{it-1}, \triangle \ln n_{it} | z_{it-1}\right) = 0 \tag{17}$$

to hold, and as $\triangle \ln \tau^p_{it-1}$ in this case is a function of x_{it-1}, the exclusion restriction corresponds to an assumption of common trends in potential income across individuals with different x_{it-1}. When estimated using within income variation, we can, in other words, think of the standard Gruber & Saez (2002) approach as a DiD estimator, and thus, validation follows the usual examination of pre-trends.

There are overall two options for such pre-trend examination. One option is to identify and split the population directly on the relevant x_{it-1}. The large literature on the effects of the earned income tax credit (EITC) (Eissa & Liebman, 1996; Meyer & Rosenbaum, 2001; Kleven, 2020) is a prominent example of this approach. Splitting the population directly on x_{it-1} is arguably the most transparent and elegant estimation strategy, but it typically requires that the tax variation is a simple function of a few variables (as, e.g., in the case of

[27] In contrast to the between variation case considered in section 2, the first stage of the standard estimation approach, $\text{cov}\left(\triangle \ln \tau^p_{it-1}, \triangle \ln \tau_{it} | z_{it-1}\right)$, is different from zero in the case with within income variation.

the EITC). However, in many cases the within income tax variation is created by changes in the treatment of various underlying income sources (e.g., capital income), specific itemized deductions, jointness in the taxation of households etc., and in these cases, the variation is likely to be too muddy to be applied in a standard DiD.

This leads to the second option of using a tax simulator to assign reform driven changes in marginal net-of-tax rates to individuals both in the actual reform period and in pre-reform periods as placebo changes (i.e., assigning marginal net-of-tax rates as if the reform was implemented in earlier periods) and compare the estimates from the actual reform to the placebo estimates from pre-reform periods. Using a tax simulator in this way is a "one-size-fits-all" approach that works irrespective of the source of the tax variation, but only by maintaining an element of black box in the estimation, where it is not clear what part of the reform variation is driven the estimates.

Within Estimates from the 2004 and 2009-10 Tax Reforms

Mirroring the between income estimates presented in Section 4 of the main paper, we estimate income elasticities using the within income tax variation created by the 2004 and 2009-10 tax reforms.

Figure A.IV shows the changes in the net-of-tax rate and within income tax variation across the income distribution for the 2004-reform (Panels A-B) and 2009-10 reform (Panels C-D). The figure complements Panel A of Figure 3 and Figure 4 in the main paper by showing the median and selected interpercentile ranges (Panels A-C) and within income tax variance (Panels B-D) of the changes in the marginal net-of-tax rates created by the two tax reforms. For the 2009-10 reform, the median changes closely follow the changes in statutory tax rates, which drive the between income variation used in the main paper, but around this level, we see a dispersion of around 10 log points for income levels below DKK 500,000 and with particular concentrated within variation between DKK 350,000-450,000. Similarly, for the 2004 reform, the median changes also closely follow the changes in the statutory tax rates, which drives the between income variation used in the main paper.[28]

[28]The dispersion of the changes in net-of-tax rates for both the 2004 and 2009-10 tax reforms is predominately created by the fact that the middle tax was based on couples' total income. Hence, differences in spousal incomes create differences in the income level at which individuals start paying the middle tax, and

Next, in Figure A.V and Table A.II we show the 2SLS elasticity estimates of the taxable income (Panels A-C) and broad income (Panels B-D). Specifically, we run

$$\Delta \ln z_{it} = \gamma_0 + \boldsymbol{\gamma_1}' \boldsymbol{D_{it}^{inc}} + \mu_2 D_{it}^{post} + \gamma_3 \Delta \ln \tau_{it} + \nu_{it}, \tag{18}$$

using only a single period (e.g., the reform period 2008-12) and where we instrument the endogenous variable $(\ln \tau_{it})$ with the within income tax variation $\left(\triangle \ln \tau_{it-1}^{p} - \triangle \ln \bar{\tau}_{zt-1}^{p}\right)$.[29] In Figure A.V, we explore heterogeneity across the income distribution by running the estimation separately for 10 equally sized groups of the population, while in Table A.II we estimate average elasticities. The exercise reveals significantly higher elasticities for primarily the lower income groups for the 2009-10 tax reform, while for the 2004 reform we find the highest elasticities for incomes around DKK 300,000.[30]

To validate these estimates we implement the placebo exercise based on the tax simulator outlined above. For the 2009-10 reform, we implement the reform in the pre-period 2004-08, where the actual tax system was close to stable. More concretely, we use our tax simulator to assign marginal tax rates to individuals in 2004 and 2008 based on the tax system in 2008 and 2012, respectively, and use as instruments the predicted changes in marginal tax rates from 2008 to 2012, while holding fixed individuals' income (and covariates) from 2004. Similarly, we implement the 2004-reform in the 1999-2003 pre-period.

For the 2009-10 reform, running the 2SLS estimator in equation (18) on the placebo period generally yields smaller elasticities (Figure A.V, Panels C-D); however, we still find positive and significant average elasticities (Table A.II, Panel B). Hence, as a whole, the groups treated differently by the within income tax variation created by the 2009-10 reform were not on parallel trends prior to the reform. Examining the 2009-10 placebo estimates across the income distribution the picture is more mixed. For initial incomes below

whether they are affected by the increase in the middle tax threshold in the 2004 reform and the abolition of the tax in the 2009-10 reform. A more tailored estimation strategy would exploit this jointness more explicitly instead of the "one-size-fits-all" approach applying in this section, which also captures tax variation created by municipal tax changes and other minor sources.

[29] In practice, we could just as well have instrumented with the total tax variation, as with only a single period and saturated income controls (DKK 10,000 bins) we soak up all between income variation.

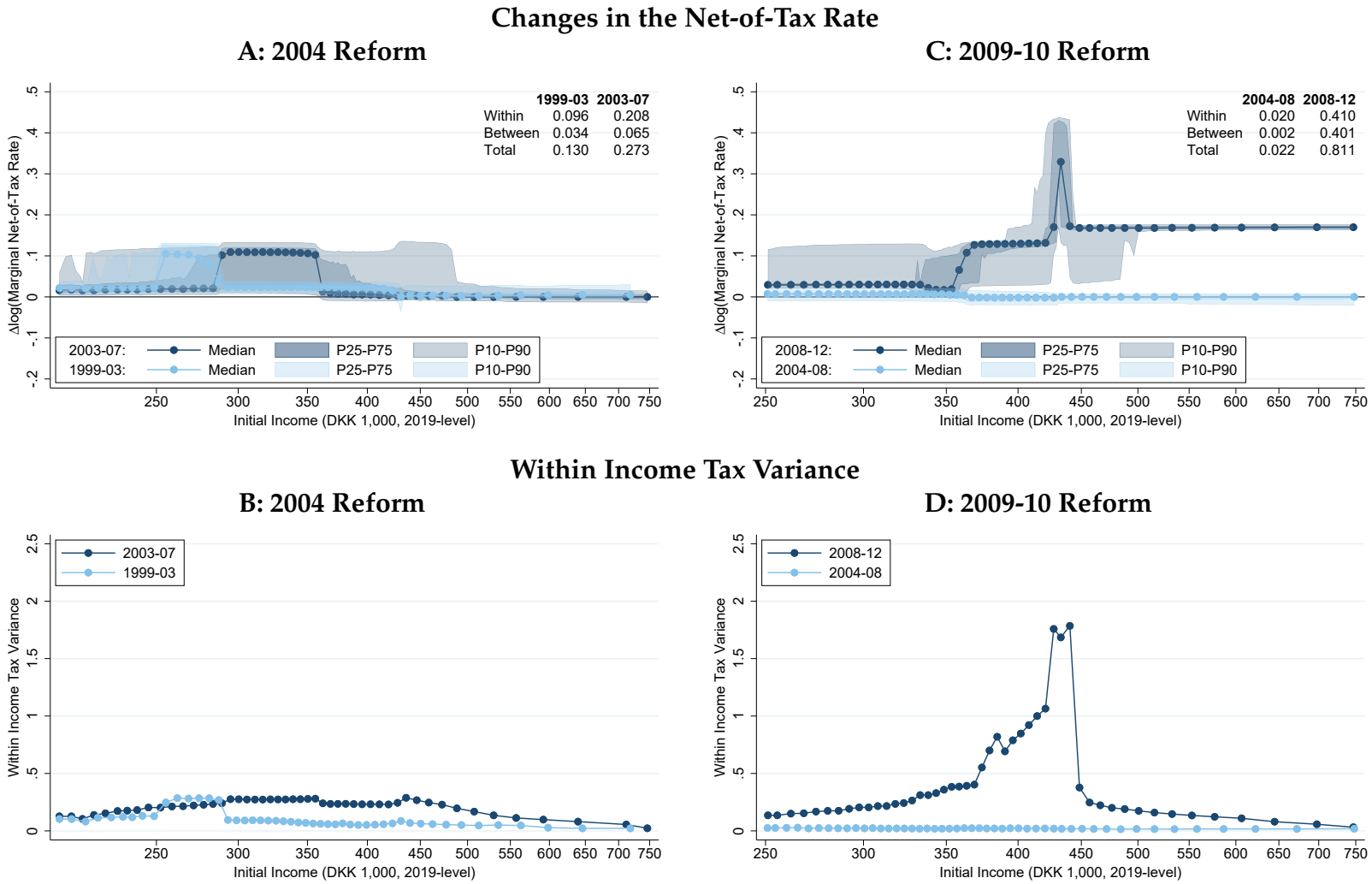
[30] The 2SLS estimator implicitly weights heterogeneous treatment effects by their relative variance in treatment intensity, and hence, for the 2009-10 reform the estimator puts more weight on the relatively low estimates for incomes between DKK 350,000-450,000 when computing the aggregate elasticity.

DKK 325,000 the estimates are close to zero and insignificant, while for incomes above DKK 450,000 we see significant placebo estimates of the same magnitude as the reform estimates.[31]

For the 2004 reform, we find an elasticity for taxable income of -0.038 for the reform period and -0.088 using the placebo period, and examining the elasticity estimates across the income distribution, we see that these negative estimates are driven by the top of the income distribution. Furthermore, we see that the reform estimates closely follow the placebo estimate except for incomes in the range of DKK 300,000-350,000.

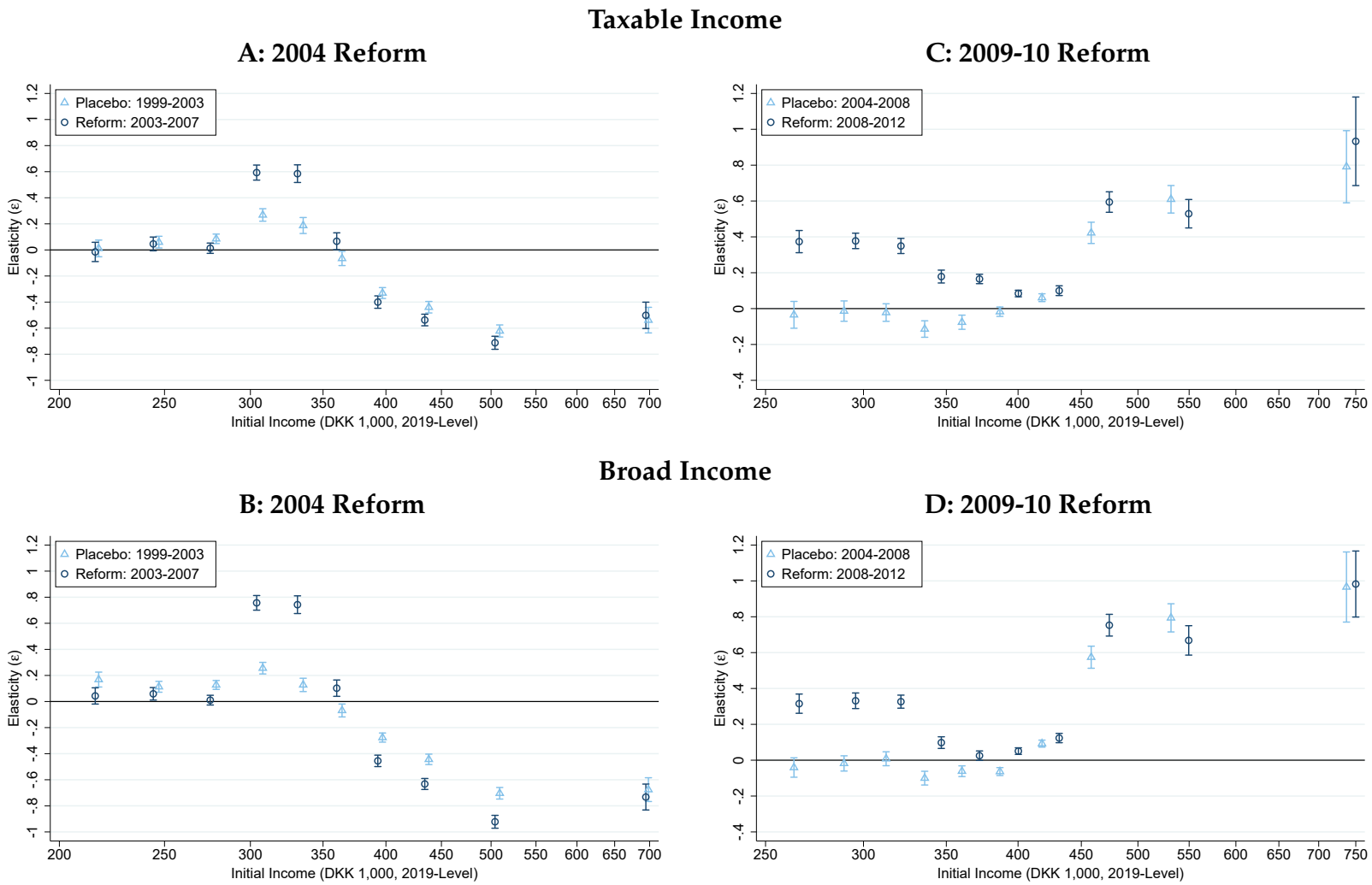
[31] In cases where the pre-period also contain variation in tax rates, the placebo exercise is only informative if the reform changes in marginal tax rates are uncorrelated with the pre-reform changes. Hence, in Table A.II we also include the results from a regression of the placebo reform changes in tax rates on the actual changes in tax rates in the pre-reform period. For the 2009-10, we only find a very small estimate due to the stability of the tax system in the pre-reform period. For the 2004 reform we find a negative estimate of -0.195.

Figure A.IV: Within Income Tax Variation Created by the 2004 and 2009-10 Reforms



Notes: The figure shows the median and selected interpercentile ranges (Panels A-C) and within income variance (Panels B-D) of the changes in the net-of-tax rate created by the 2004 and 2009-10 reforms. As in Section 2, we define within income tax variation as $\triangle \ln \tau^p_{it-1} - \triangle \ln \bar{\tau}^p_{zt-1}$, where $\triangle \ln \bar{\tau}^p_{zt-1}$ is computed as the average within DKK 1,000 bins.

Figure A.V: Elasticity Estimates Using Only Within Income Tax Variation



Notes: The figure shows the 2SLS estimates from equation (18). The point estimates in the figure are separate estimates for 10 equally sized groups of the population. For the placebo estimates, we use our tax simulator and implement the 2004 and 2009-10 tax reforms 4 years earlier (in 2000 and 2005-06, respectively) as explained in the text.

Table A.II: Income Elasticities Using Within Income Tax Variation from the 2004 Reform and the 2009-10 Reform

	Reform		**Placebo**		
Outcome	Taxable Income	Broad Income	Taxable Income	Broad Income	Actual Tax Changes
Panel A: 2004 Reform					
Second Stage					
$\triangle \log(\tau_{it})$	-0.038	-0.035	-0.088	-0.074	N/A
	(0.009)	(0.009)	(0.008)	(0.007)	
First Stage					
$\triangle \log(\tau_{it}^{P})$	0.658	0.658	0.694	0.694	-0.195
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Income Controls:	DKK 10,000 bins				
Time Periods	2003-2007	2003-2007	1999-2003	1999-2003	1999-2003
Obs.	2,295,922	2,295,920	2,240,332	2,240,330	2,240,332
Panel B: 2009-10 Reform					
Second Stage					
$\triangle \log(\tau_{it})$	0.211	0.184	0.051	0.073	N/A
	(0.006)	(0.006)	(0.007)	(0.006)	
First Stage					
$\triangle \log(\tau_{it}^{P})$	0.741	0.741	0.756	0.756	-0.019
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Income Controls:	DKK 10,000 bins				
Time Periods	2008-2012	2008-2012	2004-2008	2004-2008	2004-2008
Obs.	2,240,833	2,240,832	2,097,862	2,097,861	2,097,862

Notes: The table shows the details of the 2SLS estimates based on equation (18) for the whole sample in Figure A.V. Robust standard errors in parentheses.

D Estimating Income and Substitution Effects

In the main analysis, we disregarded the distinction between income and substitution effects of tax changes and simply sought to estimate what, in practice, is uncompensated elasticities. However, many papers in the empirical tax literature (including Gruber & Saez, 2002 and Kleven & Schultz, 2014) try to split the behavioral responses into income and substitution effects, and in this appendix, we revisit these efforts in light of our new approach presented in the main paper.

Most papers motivate the study of income effects by linearizing the (net-of) tax schedule using virtual income. In this case, let the uncompensated supply of taxable income (z) be defined as $z(\tau, \widetilde{y})$, where τ is the marginal net-of-tax rate and $\widetilde{y} \equiv z(1-\tau) - T(z) + y$ is virtual income with $T(z)$ being the tax function and y being other (untaxed) sources of income. Following the derivations in Gruber & Saez (2002), we can write the uncompensated change in z following a reform that changes τ and $\widetilde{y}$ as

$$dz = \frac{\partial z}{\partial \tau} d\tau + \frac{\partial z}{\partial \widetilde{y}} d\widetilde{y} \Leftrightarrow \frac{dz}{z} = \varepsilon_u \frac{d\tau}{\tau} + \eta \frac{d\widetilde{y}}{\widetilde{y}}, \tag{19}$$

where $\varepsilon_u = \frac{\partial z}{\partial \tau}\frac{\tau}{z}$ is the uncompensated elasticity and $\eta = \frac{\partial z}{\partial \widetilde{y}}\frac{\widetilde{y}}{z}$ is the elasticity wrt. virtual income. Further, using the Slutsky equation $\varepsilon_u = \varepsilon_c + \eta \tau z / \widetilde{y}$, we can rewrite equation (19) to

$$\frac{dz}{z} = \left(\varepsilon_c + \eta \frac{\tau z}{\widetilde{y}} \right) \frac{d\tau}{\tau} + \eta \frac{d\widetilde{y}}{\widetilde{y}} = \varepsilon_c \frac{d\tau}{\tau} + \eta \frac{dT(z)}{\widetilde{y}}, \tag{20}$$

where $dT(z)$ is the mechanical effect of the reform on the individuals' after-tax income. This equation describes the behavioral responses we should expect to see across the income distribution depending on the "structural" parameters ε_c and η.

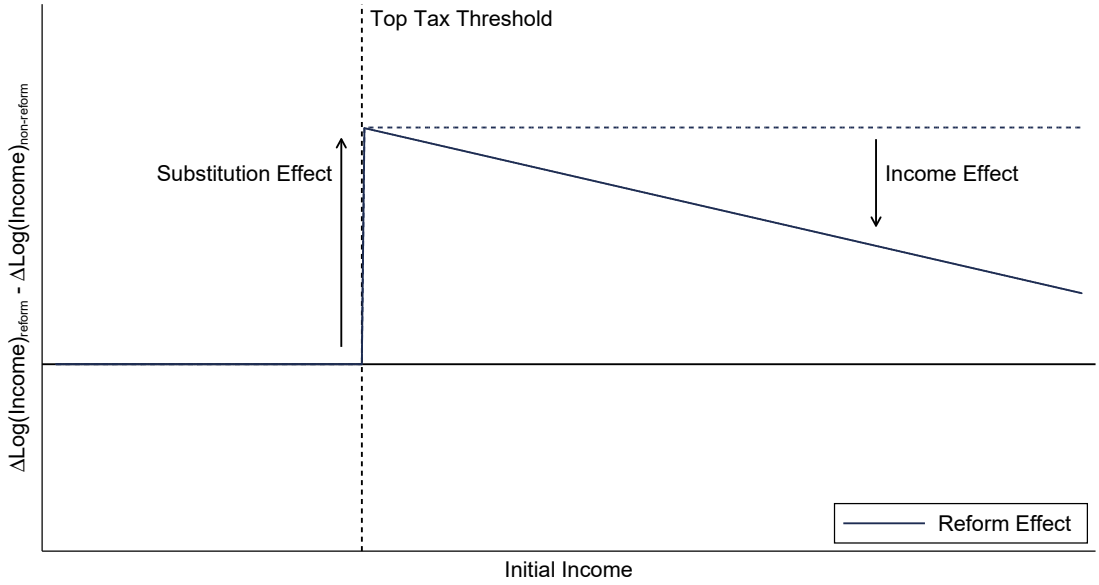
Next, let us consider a tax reform that reduces the top tax in a similar way to the reform analyzed in Section (2) of the main paper. As the reform only reduces the marginal tax rate above the top tax cutoff K, $d\tau > 0$ is constant for all top taxpayers and zero for everyone else. Thus, from equation (20) we should expect a positive and constant substitution effect for all top tax payers. In contrast, $dT(z) = d(1-\tau)(z-K)$, which is a decreasing function (as $d(1-\tau) < 0$) of the taxpayer's income that exceeds the top tax threshold. Thus, we

should expect an income effect starting at zero for tax payers just above the top tax threshold and gradually becoming more negative for higher incomes. Combining the income and substitution effects, we obtain the largest predicted behavioral responses for individuals just above the top tax threshold as these are only subject to the positive substitution effect with the effect declining for higher income due to the increasing income effect as illustrated in Figure A.VI.

Econometrically, the pattern in Figure A.VI is what we would look for in the data with the substitution effect being identified from the estimated treatment effect close to the top tax threshold, and the income effect identified from the slope of the treatment effect above. However, the presence of income effects is just one among many potential reasons for why treatment effects may vary over the income distribution. Behavioral responses to tax changes may, for example, be larger at the top of the income distribution due to income shifting, or they may be smaller close to the tax threshold due to optimization frictions such as imperfect knowledge of the exact location of the threshold, and trying to fit these heterogeneous treatment effects into a framework with a constant substitution and income elasticity is to likely yield misleading results. For example, blindly fitting the estimated elasticities from the 2009-10 reform in Figure 3 to such a framework would yield a small substitution elasticity and a large but positive income elasticity. Hence, it is not surprising that the income effects estimated in the empirical tax literature do not align with the estimates using other sources of variation, such as, for example, lottery winnings (see, e.g., Imbens *et al.*, 2001 and Cesarini *et al.*, 2017).

Other tax reforms, such as changes in tax thresholds may create variation that is better suited to separate income and substitution effects. However, estimating the heterogeneity of treatment effects non-parametrically, as we do in Figure 3, preserves the transparency of the estimation and avoids functional form assumptions of constant substitution and income elasticities.

Figure A.VI: Identifying Income and Substitution Effects



Notes: The figure illustrates the predicted behavioral responses from a reduction in the top tax rate based on equation (20), assuming constant substitution (ε_c) and income (η) elasticities.

E Examining Additional Tax and Benefit Reforms

In this appendix, we apply our new approach to two additional tax reforms in Denmark.[32] Figure 2 in the main paper shows the marginal tax rates on labor income over time, and in addition to the 2009-10 and 2004 tax reform studied in the main paper, we highlighted the reform of 1987and the reform of 1994. We analyze the 1987 and 1994 reforms in four key graphs and estimate the income elasticity using the standard approach in a similar way to our analysis of the 2004 and 2009-10 reforms in the main paper.

The 1987 Tax Reform

The 1987 tax reform changed the Danish tax system in multiple dimensions, most notably by introducing a dual tax system with lower tax rates on capital income (see also **?**). In addition, the reform reduced the marginal tax rates on labor income for tax payers in top and middle tax brackets, while increasing them slightly in the bottom bracket as illustrated in Figure 2 in the main paper. Hence, at first glance the reform variation resembles the variation created by the 2009-10 reform, and similar to our analysis in the main paper of the 2009-10 and 2004 reform, we analyze the 1987 reform in four key graphs in Figure A.VII.

We start in Panel A by showing the distribution of changes in the marginal net-of-tax rates across the income distribution in the pre-reform (1982-1986) and reform (1986-1990) periods.[33] From this panel, we notice two features that make the 1987 reform less ideal. First, the pre-period also contains considerable tax variation, and second, while both the within and between income tax variation is considerably larger in the reform period, the between income variation is more dispersed across the income distribution with less clear untreated groups. To make progress, we define the validation region as taxable incomes below DKK 260,000, as this region experiences somewhat similar changes in marginal net-of-tax rates in the two periods. Consequently, we define the identification region as taxable

[32] An overview of the Danish tax reforms since 1980 can be found on the homepage of the Ministry of Taxation (in Danish)

[33] For the 1987 reform we use legal taxable income as defined in Table 1 in the main paper as initial income. Taxable income was the main tax base for labor income in our two initial years (1982 and 1986). Personal income only become the main tax base for labor income after the reform.

income above DKK 260,000.

Panel B plots the differences in income growth across the income distribution with the change for incomes around DKK 225,000 normalized to zero. Despite the less ideal tax variation, the graph reveals the expected downward sloping pattern, which is consistent with mean reversion. In the validation region, the observed income trends largely follow the same pattern in the reform period as in the pre-reform period, while for incomes above DKK 260,000 we observe significantly higher income growth compared to the pre-reform period. The same is visible for taxable income in Panel C, which shows the estimated changes in income trend differentials from 1982-86 to 1986-90. Looking at broad income, we observe the reverse, with significantly negative changes in income trends in the identification region.

In Panel D, we translate the estimated trend differentials in Panel C into elasticities using a local linear 2SLS estimator similar to the one used in Section 4 in the main paper. In contrast to the 2009-10 reform, we find more homogenous elasticities for taxable income across the income distribution, but the implied elasticities for broad income are negative.

The 1994 Tax Reform

The 1994 reform followed the same direction as the 1987 reform by reducing marginal tax rates for labor income and widening tax bases. From Panel A in Figure A.VIII, we see that, on average, the reform increased the marginal net-of-tax rate by around 12 log points relative to the pre-reform period for incomes above DKK 325,000, while the changes in marginal tax rates for income below were similar in the reform and pre-reform periods (keeping in mind the minor differences in treatment for individuals with income below DKK 300,000).[34] Hence, we define the validation region as personal income between DKK 250,000 and DKK 325,000, and the identification region as above DKK 350,000.

Panel B plots the income trend differentials across the income distribution with the change for incomes around DKK 300,000 normalized to zero. The graph reveals the expected downward sloping pattern consistent with mean reversion, but we find no evi-

[34] It is also worth mentioning that the changes in marginal tax rates were phased in over a number of years, as illustrated in Figure 2 in the main paper.

dence of an upward shift in income growth for individuals in the identification region. In fact, we find the opposite with the 1993-97 curve lying below the pre-reform 1989-03 curve. The same is visible in Panel C, which shows the changes in income trend differentials from 1989-93 to 1993-97. In this panel, we do find higher growth for broad income in the identification region, but only for individuals with relatively high initial income. In Panel D, we translate the estimated trend differentials in Panel C into elasticities using a local linear 2SLS estimator similar to the one used in Section 4 in the main paper. Consistent with Panel C, we find heterogenous and negative estimates for taxable income and heterogenous and positive for broad income.

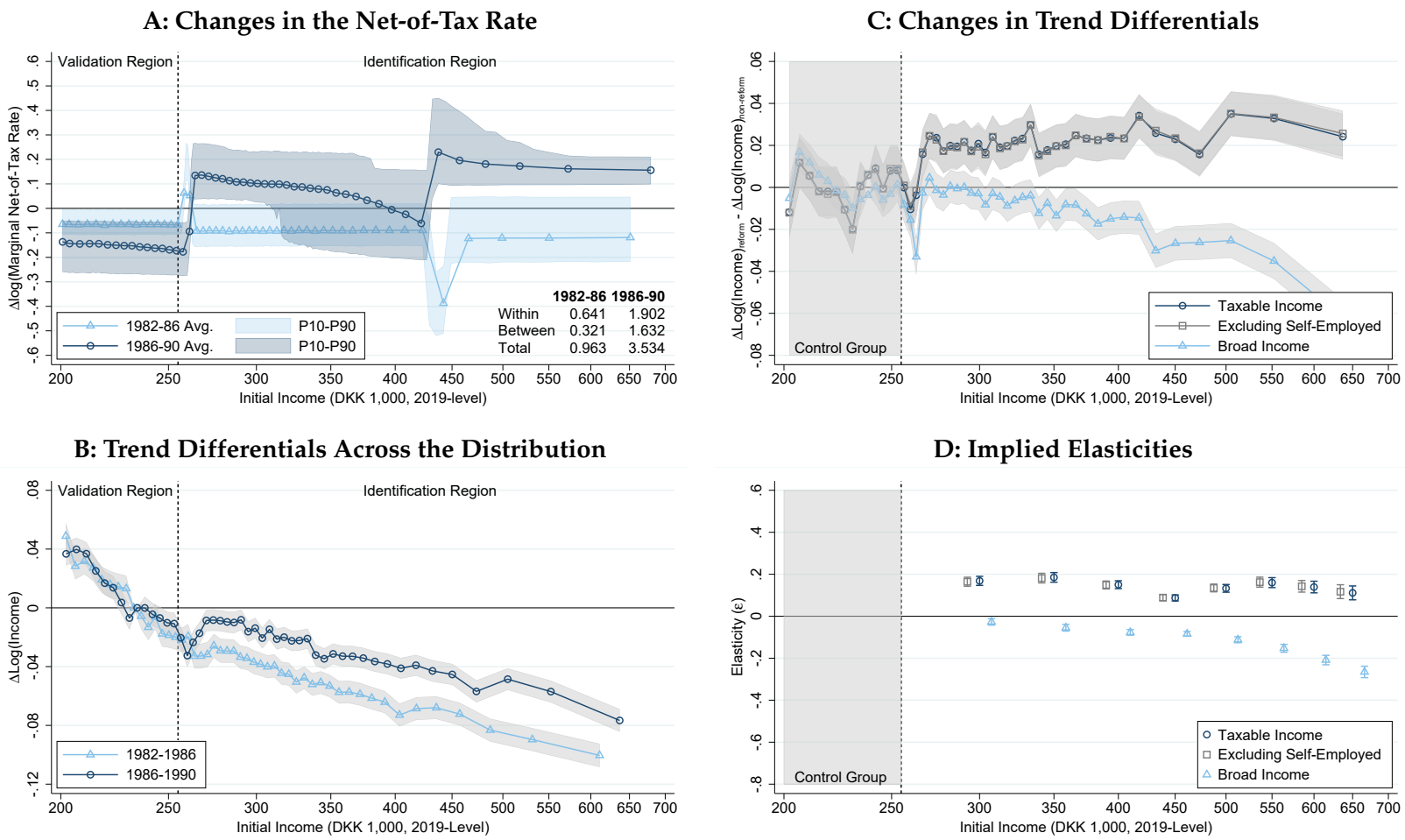
Overall, Panels B and C in Figure A.VIII raise doubts about the validity of the identifying assumption of constant trend differentials in the case of the 1994 reform, as we see significant changes in trend differentials and in the validation region, and the overall pattern of changes in Panel C appears unrelated to the tax variation.

Results From the Standard Approach

Above, we have analyzed the 1987 reform and 1994 reform using our new approach to validate the identifying assumptions by graphically inspecting the changes in trend differentials from the pre-reform to the reform periods. Parallel to the main paper, we also examine the 1987 reform and 1994 reform using the standard approach by running equation (14) separately for the two reforms and present the results in Table A.III.

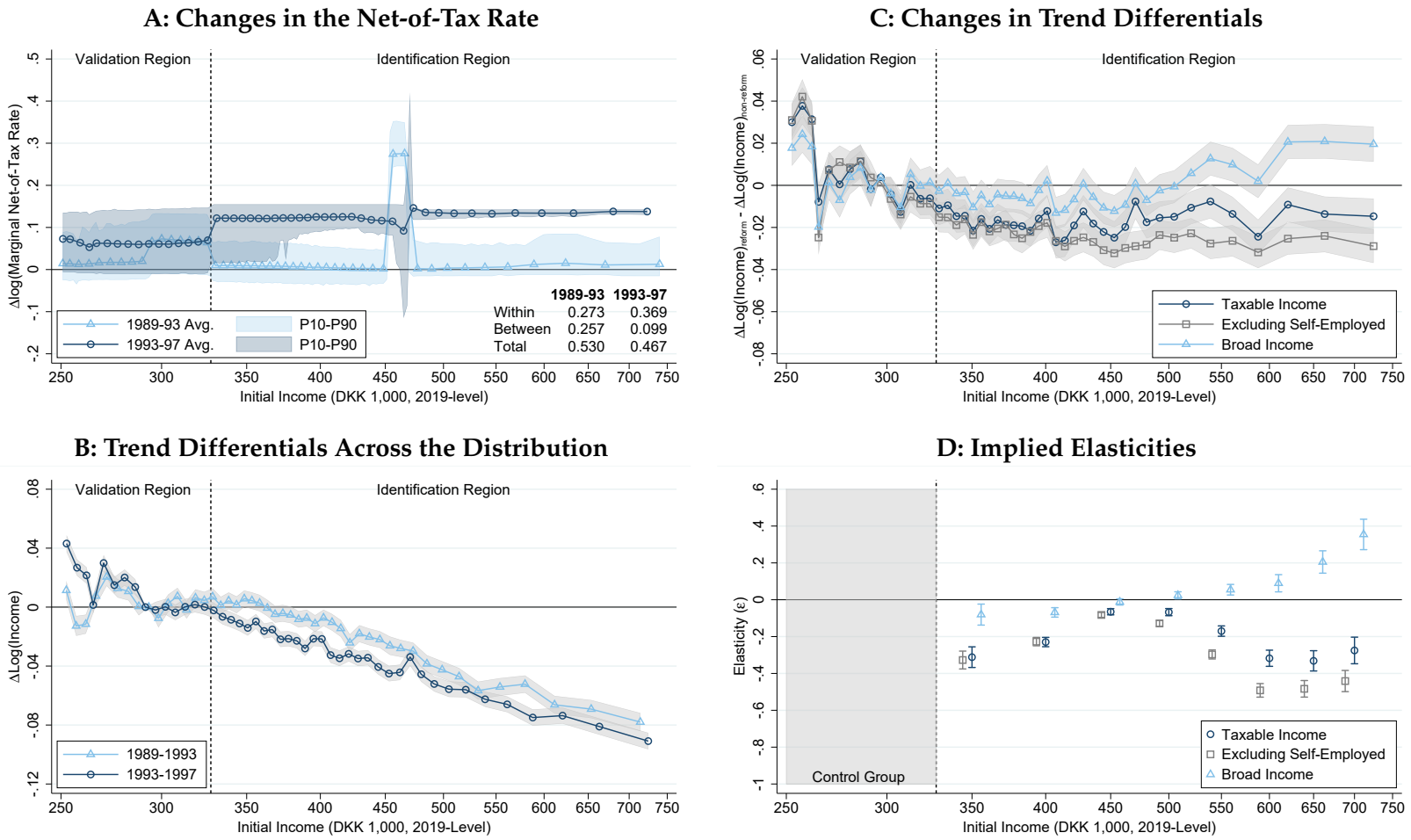
Considering first the results from the 1987 reform in columns (1)-(3), we estimate an average elasticity of 0.094 for taxable income (column 1) with and without the self-employed (column 2). Considering a broader income measure, we find a negative average elasticity (column 3). Turning to the 1994 reform in columns (4)-(6), we find a negative estimate for taxable income with (column 4) and without (column 5) the self-employed. Considering a broad income measure, we find a positive elasticity estimate. However, these estimates are less likely to be causal given our inspection of the trend differentials above.

Figure A.VII: Identifying Income Responses Using the 1987 Tax Reform



Notes: Panel A shows the predicted changes in the log net-of-tax rate (τ^p_{it-1}) from 1982 to 1986 (the pre-reform period) and 1986 to 1990 (the reform period). The curves show the average changes within each income bin and the shaded areas shows the 10th and 90th percentile ranges. Panel B shows the estimated income trend differentials using equation (10) for 1982-86 and 1986-90 relative to the growth rate for income around DKK 260,000. Panel C shows the estimated changes in trend differentials based on equation (11) for different samples and income concepts. The use of different income concepts only affects the dependent variable (y-axis). We always use taxable income as initial income (x-axis) as taxable income most accurately determines treated by the reform. Initial income is measured in 1982 and 1986 and thus is pre-reform. Panel D shows the implied elasticities over the income distribution estimated using the 2SLS local linear estimation described in equation (12). USD 1 $\approx$ DKK 6.8.

Figure A.VIII: Identifying Income Responses Using the 1994 Tax Reform



Notes: Panel A shows the predicted changes in the log net-of-tax rate (τ^p_{it-1}) from 1989 to 1993 (the pre-reform period) and 1993 to 1997 (the reform period). The curves show the average changes within each income bin and the shaded areas shows the 10th and 90th percentile ranges. Panel B shows the estimated income trend differentials using equation (10) for 1989-93 and 1993-97 relative to the growth rate for income around DKK 300,000. Panel C shows the estimated changes in trend differentials based on equation (11) for different samples and income concepts. The use of different income concepts only affects the dependent variable (y-axis). We always use taxable income as initial income (x-axis) as taxable income most accurately determines treated by the reform. Initial income is measured in 1989 and 1993 and thus is pre-reform. Panel D shows the implied elasticities over the income distribution estimated using the 2SLS local linear estimation described in equation (12). USD 1 $\approx$ DKK 6.8.

Table A.III: Standard Approach Income Elasticity Estimates from the 1994 Reform and 1987 Reform

	1987 Reform			**1994 Reform**		
	Taxable Income	Excluding Self-Employed	Broad Income	Taxable Income	Excluding Self-Employed	Broad Income
	(1)	(2)	(3)	(4)	(5)	(6)
Second Stage						
$\triangle \log(\tau_{it})$	0.094	0.094	-0.089	-0.056	-0.094	0.011
	(0.006)	(0.006)	(0.005)	(0.008)	(0.007)	(0.007)
First Stage						
$\triangle \log(\tau_{it}^{P})$	0.522	0.523	0.523	0.718	0.738	0.738
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Income Controls	DKK 10,000 Bins					
Time Periods	1982-1986	1982-1986	1982-1986	1989-1993	1989-1993	1989-1993
	1986-1990	1986-1990	1986-1990	1993-1997	1993-1997	1993-1997
Obs.	3,672,381	3,635,368	3,635,368	4,111,030	3,603,492	3,603,492

Notes: The table summarizes the results from the standard estimation approach described in equation (14) for the 1987 and 1994 reforms. Robust standard errors in parentheses.