Economics
Working Papers

2022-04

Labor Supply Responsiveness to Tax Reforms

Hans Schytte Sigaard

AARHUS BSS
DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

Labor Supply Responsiveness to Tax Reforms*

Hans Schytte Sigaard

Department of Economics and Business Economics, Aarhus University

August 30, 2022

Abstract

Labor supply responses constitute real responses to taxation and are central for policy analysis. This paper estimates the elasticity of labor supply at the intensive margin by applying conventional estimation strategies from the elasticity of taxable income (ETI) literature to administrative register data on performed hours of work. The elasticity is estimated to be 0.08, significant, and can be attributed to individuals changing main and secondary employment, a general increase in contract hours, and a reduction in paid absence. By also estimating ETI, which captures all responses to taxation, I show that a large part of total responses can be attributed to individuals adjusting work hours, capturing inherent labor responses, and I thereby provide a link between the labor and public finance literatures. These findings suggest that contextual factors such as labor market fluidity and flexibility are important to facilitate real responses to taxation and that labor supply responses are core underlying drivers of total responses to taxation.

Keywords: Elasticity of labor supply, marginal tax rates, tax reforms, behavioral responses, public economics

JEL codes: H21, H24, H30, J22, J24, J62

*I gratefully acknowledge comments and suggestions from Rune Vejlin, Michael Svarer, Jakob Søgaard, Håkan Selin, Aarhus University seminar participants, and participants at Spring Meeting of Young Economists 2022. And a special thanks to the Danish Ministry of Finance for providing codes used to simulate marginal taxes. All remaining errors are my own.

1 Introduction

Individuals may respond to changes in taxation along several margins, including changes in work effort, labor supply, job location, and tax avoidance. The magnitude of these behavioral responses are crucial for economic policy and to evaluate efficiency costs of taxation. Under certain conditions, the elasticity of taxable income (ETI) has been shown to capture all relevant responses for efficiency and optimal taxation (Feldstein, 1995, 1999) as a sum. Using the hierarchy of Slemrod (1990, 1995), the ETI contains three tiers of behavioral responses: real, avoidance, and timing responses. In a narrow perspective, where the tax system is given, the relative sizes of these responses are irrelevant for e.g. the government budget. In a broader perspective, where it is possible to design the tax system to reduce avoidance opportunities and increase tax enforcement, the type of behavioral responses become crucial (Saez et al., 2012). If avoidance and timing responses can be effectively minimized, real responses are what remains. Responses in labor supply, in the form of leisure choices and firms' inputs in production functions, are at the core of the real responses in Slemrod's hierarchy. Therefore, estimating the responsiveness of labor supply is vital for economic policy decisions.

There are several papers documenting large avoidance and timing responses[1], and it is the standard view among many public economists that the ETI is primarily driven by such responses (Slemrod, 1995; Slemrod and Yitzhaki, 2002; Saez et al., 2012; Harju and Matikka, 2016). Contrary to this view, I show, by estimating the elasticity of labor supply (ELS) along with ETI, that adjustments to hours of work is at the core of real responses and the primary driver of total responses for wage earners in Denmark. I argue that this is because of two reasons. Firstly, avoidance and timing responses are low for wage earners who are subject to third-party reporting in Denmark (Kleven et al., 2011; Kleven and Schultz, 2014), and secondly, the response in hours of work is large due to the fluidity and flexibility of the Danish labor market. I show that job switches and adjustments to paid absence in particular make hours of work responsive in Denmark. Linking the ETI and ELS literatures by estimating ELS in addition to ETI provides a ratio of real to total responses. The analysis complements the ETI literature by decomposing the total effect and identifying potential channels for the real responses, which is necessary for calculating the efficiency cost of taxation (Chetty, 2009).

Estimating the ratio of real to total responses by comparing ELS to ETI is of high interest, yet has been difficult for two reasons. Firstly, large discrepancies in methodological approaches make comparison difficult. Secondly, a practical difficulty in using hours of work as the outcome of interest is the

[1] For avoidance responses, see e.g. Kreiner et al. (2014); Fack and Landais (2016); Doerrenberg et al. (2017) or Neisser (2021) and for timing responses e.g. Goolsbee (2000) or Kreiner et al. (2016).

availability and precision of such a measure. Previous labor supply literature (see Section 1.1) has relied on survey data on hours of work, which are imprecise, have limited coverage of the population, and often consist of short individual panels. It is unlikely that full behavioral responses have materialized within a single year, and it is standard to account for slow, "sluggish", responses that take several years to materialize in the ETI literature. In this paper, I use "performed" work hours from large and detailed administrative register data. Register data include longer individual panels, which makes it possible to calculate individual differences in marginal taxes and hours of work over time, to account for sluggish adjustments. The advantage of using *performed* hours of work is that it includes paid overtime[2] and excludes paid absence (e.g. due to sickness, maternity leave, special holidays and similar). In this way, it captures variation in contracted hours of work and to some extent, variation in effort, as measured by taking more paid overtime and/or less paid leave. The increasing availability of a precise and detailed measure of hours of work is promising for future estimation of labor supply responses.

Using this data, I make three key contributions to previous literature. Firstly, I estimate the three-year ELS and ETI to be 0.08 and 0.1 respectively, and thereby provide a relation between behavioral responses in labor supply (ELS) and total behavioral responses to tax rate changes (ETI). Labor supply responses is a highly appealing outcome variable because it reflects real responses, not subject to intertemporal shifting or tax avoidance. In this way, I link the ETI and ELS literatures and is, to my knowledge, the first to provide evidence that a large part of the total behavioral responses stem from real responses by individuals adjusting their hours of work. The analysis should be considered a complement to the ETI literature by documenting real responses as a core underlying driver of the total effect of taxation.

My second contribution is by decomposing labor supply responses into within-job and across-job variation. One may mistakenly have the impression that labor supply analyses rely on the assumption that individuals can freely choose their work hours at a fixed wage with a single employer (Blundell and Macurdy, 1999). Instead, individual variation in hours of work can come from two different sources; within-job or across-job variation. Within-job variation can come from individuals renegotiating their job contract to obtain more hours (e.g. going from part-time to full-time work), taking more overtime work or making less use of paid absence. Across-job variation can come from the individual switching to a job with more work hours or taking up new side employment. Blundell and Macurdy (1999) interpret labor market models as characterizing situations where individuals select their hours of work by selecting across employers offering different wage packages. I find evidence that the estimated ELS is primarily

[2] Although *unpaid* overtime is unavailable.

driven by individuals switching main and secondary employment, by increasing their contract hours, and decreasing paid absence. I also find a smaller, positive effect on paid overtime. Across-job variation being a dominant channel implies that ELS will be larger, the easier job-to-job transitions are. An implication is that ELS depends on labor market conditions and tightness.

The third contribution is methodological, showing that marginal tax instruments based on lagged income information as suggested by Weber (2014) (see Section 1.1) is not always feasible and comes with a cost. Basing instruments on further lagged income reduces their explanatory powers, which may even increase asymptotic bias. Additionally, instruments based on further lagged income changes the compliant population. This may lead to a compliant population which is characterized by a higher/lower responsiveness to tax rate changes, which increases/decreases the estimated elasticity and might lower external validity.

These three contributions are largely made by applying the methodological advancements made in the ETI literature to estimating ELS, using large and detailed administrative register data. Even if ETI fulfills the conditions for being a sufficient statistic for welfare analysis, all three tiers of responses in Slemrod's hierarchy are tangled together. The primary objective of this paper is to return to labor supply responses as the outcome of interest, which only reflects real responses, and thereby disentangle labor responses from the total effect as estimated by ETI.

1.1 Relation to Previous Literature

The ELS has been notoriously difficult to estimate. The topic is plagued by endogeneity and sensitivity to modeling approach, causing a great variation in previous study results and an equally large variation in estimation approaches. The difficulties researchers have typically faced include non-linear tax-schedules, differing individuals' tastes over time that cannot be controlled for, as well as the observed individuals' decisions reflecting both inter-temporal and within-period allocations (Blundell et al., 1998).

More recently, focus in the literature has switched from estimating the ELS to ETI. The main advantage of this concept is it includes all taxpayers' responses to tax rate changes, all of which reflect inefficiencies, as these responses would not be made, absent of taxes. The ETI is used as a sufficient statistic for welfare analysis (Feldstein, 1999; Chetty, 2009), meaning one does not need to inquire into the anatomy of behavioral responses. Although, this sufficient statistic property tends to be overstated (Saez et al., 2012). For instance, given the purpose of most deductions, they typically generate non-negligible externalities (e.g. charitable contributions), the result being failing to meet the conditions for being a sufficient statistic (Doerrenberg et al., 2017). In addition to changes in labor supply, the ETI will capture changes in work-effort, tax evasion and extent of undeclared work. Contrarily, the ETI also captures effects that are irrelevant, e.g. for calculating the degree of self-financing for a tax reform. In particular, intertemporal shifting of income can play a large role for calculating ETI (as shown by e.g. Kreiner et al. (2016)).

The empirical analysis used in this paper follows the conventional ETI literature. Inspiration is drawn from both Kleven and Schultz (2014) and Weber (2014), whom in turn follow Gruber and Saez (2002). The conventional ETI approach models taxable income as a function of the marginal tax rate. To account for reverse causality, the standard strategy is to use exogenous variation induced by tax reforms by predicting changes in marginal tax rates using only base-year ("pre-reform") income. While variation induced by tax reforms is exogenous to post-reform incomes, it still depends on pre-reform income levels, which is correlated with the error term through mean reversion and heterogeneous growth rates in income. This is commonly conceptualized by modeling income as an autoregressive process, where endogeneity arises due to serial correlation. Gruber and Saez (2002) and Kleven and Schultz (2014) control for the endogeneity by including (non-linear) controls for base-year income. Income controls used in this manner can also be interpreted as controlling for non-parallel trends in a difference-in-differences setting. Weber (2014) shows that base-year income controls are endogenous under a serially correlated error term of the income process, but that endogeneity is reduced arbitrarily much by basing instruments and income controls on further lagged income. The previously cited studies all primarily exploit tax rate changes that vary across income distribution. In a recent study by Kumar and Liang (2020), the authors argue that using the conventional variation in marginal taxes caused by variation across income distributions is problematic because income is a function of ETI. Instead, they characterize changes in budget sets by a vector of marginal tax rate changes at various points in the income distribution. In this way, they use variation across taxpayers, in contrast to income distributions. Kumar and Liang (2020) argue that discrepancies in previous study results can be explained by differences in weights, determined by degree of compliance to tax rate changes, given to different taxpayers.

The standard ETI approach with overlapping differences has received critique for systematically underestimating the true elasticities if the behavioral adjustment process continues beyond the year of implementation. This is shown by papers such as Holmlund and Söderström (2008); Bækgaard (2014), and Vattø (2020). They instead estimate short-run and medium-run effects separately by estimating dynamic income models. As an alternative to the conventional ETI methods, Jakobsen and Søgaard (2022) propose a graphic validation approach, similar to validation of pre-trends in difference-in-differences studies. The main advantage of their approach is the higher transparency in validation of the identifying assumptions and visual appeal.

The previous empirical work using hours of work as the outcome variable is in general older, as focus has shifted to estimating the ETI. Frederiksen et al. (2001) estimate a "Hausman labor supply model" to explicitly treat non-linearities imposed by the tax system. The Hausman model is, however, also known to be sensitive to specification of functional form and measurement errors, and very demanding in terms of information required to construct the budget constraint (Blundell and Macurdy, 1999). Bingley and Lanot (2002) estimate the elasticity of labor supply in Denmark by using a static model for equilibrium wages and labor supplies. Blundell et al. (1998) estimate labor supply elasticities by using grouping estimators based on birth cohort and education level. Burns and Ziliak (2017) (estimating ETI) draw inspiration from the grouping estimators approach by Blundell et al. (1998). In a similar vein, Meyer and Rosenbaum (2001) and Gelber and Mitchell (2012) impute income based on demographics and calculate marginal taxes based on these income levels to model the extensive margin of labor supply.

While the outcome variable in this paper is hours of work, as in the ELS literature, I apply the methodology proposed in the conventional ETI literature. Thereby, keeping a close connection between the two strands of literature.

2 Taxation of Labor Income in Denmark

Different forms of income are taxed in a partially separated fashion in Denmark. In general, labor income faces a higher tax than capital income and is the main source of income for most individuals. The primary forms of income consist of labor income (LI), personal income (PI, labor income and other personal income minus labor market contributions and certain pension contributions), capital income (CI, excluding stock income), and deductions (D). Taxable income (TI), is given by the sum of personal income and capital income net of deductions (TI = PI + CI - D), according to the Danish tax code. Stock income (SI) is also taxable income, but belongs to a separate tax schedule and is accordingly omitted from the previous definition. Broad income (BI) is defined as labor income plus capital income (BI = LI + CI).

There are three tax brackets: bottom, middle and top. The tax rates given by these are cumulative, meaning that a taxpayer in the top-bracket pays bottom, middle and top taxes, along with all other relevant flat taxes. The middle-bracket tax was abandoned in 2010. "Labor market contributions" is a flat 8 % (for 1997 and after) tax on all labor income before any deductibles are applied. Deductibles relevant for labor income include personal, employment, and job allowances as well as deductions on interest payments. Bottom-bracket tax is applied to personal income that is above the personal allowance deduction, while municipal tax, health contributions, and church tax are applied to the taxable income exceeding the personal allowance among other income deductions. Municipal tax varies by individual, as it depends on the person's municipality of residence. Each municipality individually determines the tax its citizens must pay. Top-bracket tax is finally applied to personal income that exceeds the top-bracket income boundary. There is a maximum tax rate an individual can achieve, denoted by the "tax ceiling". The tax ceiling dictates that the sum of the other taxes cannot exceed a certain level (52.06 % in 2020). If the tax rate were to exceed this ceiling, a reduction in the top-bracket tax rate will be applied.

During the period 1999-2016, there have been 5 tax reforms in Denmark. The most important features are summarized by Table 1. In general, the reforms have targeted the tax bracket thresholds and deductions, with relatively few changes to tax rates. The most influential change was the abolition of the middle-bracket tax in 2010. Figure A.1a illustrates the changes made to the middle- and top-bracket thresholds over time (in current DKK). The top-bracket increased relatively steadily until the 2009 and 2012 reforms, where the threshold increased in both cases. From this figure, the 2004, 2007, and 2009 reform changes targeting the middle-bracket threshold are also clear. In Figure A.2a, the average marginal tax by each tax-bracket is illustrated. Note, since the middle-bracket threshold coincided with the top-bracket in 2009, all middle taxpayers were also top taxpayers. From this figure, it is very apparent that the 2009 tax-reform was highly effective in reducing the marginal tax for top taxpayers. In general, there has been steady decline in marginal taxes for all brackets throughout the period. Figure A.2b illustrates the share of the population in the middle-bracket declining over time until its abolition in 2010. The share of the population in the bottom-bracket has increased sharply over time. The share in the top-bracket has remained relatively stable, with the exception of the 2009 and 2012 reforms leading to relatively large declines. In addition to the national changes summarized by Table 1, changes in municipality taxes provide regional variation in marginal taxes. These changes are also exploited in the following empirical analysis, but are generally smaller and more uniform across taxpayers within regions.

Table 1: Overview of Tax Reforms 1999-2016

Reform year	Main points
1999	• Higher gradual increases to middle- and top-bracket thresholds • Bottom-bracket tax rate decreased (1999-02 from 7.5 to 5.5 %)
2003	• Employment deduction introduced (2004 - max 7,000 DKK) • Middle-bracket threshold increased (2003-04 from 198,000 to 254,000 DKK)
2007	• Middle-bracket threshold increased to coincide with top-tax bracket (2008-09 from 279,800 to 347,200 DKK) • Health contributions (8 %) introduced to partially replace county taxes
2009	• Middle-bracket tax abandoned (2010) • Bottom-bracket tax rate decreased (2009-10 from 5.04 to 3.67 %) • Increase in top-bracket threshold (2009-10 from 347,200 to 389,900 DKK)
2012	• Top-bracket threshold increased (2012-13 from 389,900 to 421,000 DKK with gradual increases until 2023)

This table contains an overview of the main points of the Danish tax reforms implemented in the period 1999-2016. The column "Reform year" displays when the tax laws were passed in parliament. Most changes were implemented in the following year, as noted in the column "Main points". All monetary values are nominal. 1 USD ≈ 6.2 DKK, as of August 31, 2020.
Sources: www.skm.dk (1), www.skm.dk (2), www.dors.dk

3 Empirical Model Setup

The theoretical foundation for labor supply modeling is based on a quasi-concave utility function and a budget constraint.[3] The consequent utility maximization problem can be solved to derive the individual Marshallian demand for work hours. Differentiating the log of demand for work hours w.r.t. log of the post-tax wage rate yields the uncompensated (Marshallian) wage elasticities. "Uncompensated" refers to the elasticity including both substitution and income effects, in contrast to the compensated (Hicksian) wage elasticity, which only includes the substitution effect. The link between uncompensated and compensated elasticities is given by the Slutsky equation.[4]

As an increase in the post-tax wage is equivalent to a decrease in the marginal tax rate, I adopt the

[3]The following is only a brief summary. For an elaborate coverage, see Blundell and Macurdy (1999).

[4]The following empirical approach estimates the uncompensated elasticity and focus is neither put on the compensated elasticities nor the discrepancy between substitution and income effects.

following log-linear specification to empirically estimate the uncompensated elasticity of labor supply:

$$\Delta_s \ln h_{it} = \alpha_i + \theta_t + \varepsilon \Delta_s \ln(1-\tau_{it}) + f(\ln z_{it-k}) + X_{it}\beta + \nu_{it}, \quad (1)$$

where "Δ_s" denotes the s-year difference of the variable, in the main specification being three-year differences. Note, I define $\Delta_s \ln(1-\tau_{it}) \equiv \ln(1-\tau_{it+s}) - \ln(1-\tau_{it})$, such that year t denotes the base-year. α_i are individual fixed effects, θ_t are time fixed effects, X_{it} denotes a set of control variables elaborated below, $f(\ln z_{it-k})$ is a flexible control function for log labor income in period *t-k*, and ν_{it} is the idiosyncratic error term. $\ln(1-\tau_{it})$ is the net of marginal tax rate on additional labor income faced by individual i at time t, the associated coefficient, ε, being of main interest. ε is the uncompensated elasticity of labor supply w.r.t. the post-tax wage rate. The set of control variables, X_{it}, include individual, municipal, and industry controls, all these simply being referred to as "socio-economic controls". These include experience, experience squared, municipal dummies, unemployment rate in the municipality at time t, dummy variables for; age, being married, having a child in age ranges 0-2, 3-6, 7-9, 10-14, 15-17, education being less than high school, high school or college graduate and 10 industry groups[5] of primary workplace.

3.1 Identification

The marginal tax rate is endogenous to hours of work because of non-linearities in the tax system. This is because a higher number of hours worked increases labor income and therefore potentially the marginal tax the individual faces. This creates correlation between $\Delta_s \ln(1-\tau_{it})$ and the error term, ν_{it}. The usual way to deal with this endogeneity is by constructing predicted tax rate instruments, resulting from changes in the tax system over time (Auten and Carroll, 1999; Gruber and Saez, 2002; Kopczuk, 2005; Kleven and Schultz, 2014). That is, predicting marginal taxes at time *t+s*, using income information at time t, $\Delta_s \ln(1-\tau_{it}^p)$. These instruments are exogenous to post-reform (period *t+s*) income, but still depend on base year (period t) income. This causes endogeneity, as base year income is also correlated with the error term. The ETI literature describes the correlation coming from two primary channels: mean reversion and heterogeneous income (secular) trends. Mean reversion describes how individuals with high or low income in a period, relative to their permanent income level, will experience a return to their mean income in the following period. This is commonly interpreted as caused by the error term containing transitory income shocks, leading to it being serially correlated (see Moffitt and Wilhelm

[5] Industry groups follow the standard groupings of www.dst.dk.

(2000); Kopczuk (2005); Weber (2014)). The standard way to control for both these problems has been to include controls for base year income, but as shown by Weber (2014), inclusion of any of the previously proposed income controls will not be enough to make $\Delta_s \ln(1-\tau^p_{it})$ uncorrelated with the error term. The two problems of mean reversion and heterogeneous growth rates are covered in turn below.

3.1.1 Mean Reversion

To account for mean reversion, Weber (2014) proposes instruments using predicted changes in net of marginal tax rates using income information lagged 2-4 years. I define these for income information lagged k periods as:

$$\Delta_s \ln(1-\tau^p_{it,t-k}) \equiv \ln(1-T'_{t+s}(z_{it-k})) - \ln(1-T'_t(z_{it-k})), \tag{2}$$

where $T'_{t+s}(z_{it-k})$ is the marginal tax rate under tax system in period t+s, calculated using income information in period t-k. As in Weber (2014), the asymptotic bias of the elasticity estimate using the instruments in Equation (2) is given by the following equation:

$$\text{plim}(\hat{\varepsilon}_{IV}) = \varepsilon + \frac{\text{cov}[\Delta_s \ln(1-\tau^p_{it,t-k}), v_{it}|\chi_{it}]}{\text{cov}[\Delta_s \ln(1-\tau^p_{it,t-k}), \Delta_s \ln(1-\tau_{it})|\chi_{it}]}, \tag{3}$$

where χ_{it} denote all conditioning variables. From Equation (3), the two identifying assumptions for the IV approach are very apparent: *validity* ($|\text{cov}[\Delta_s \ln(1-\tau^p_{it,t-k}), v_{it}|\chi_{it}]|$ is small) and *relevancy* ($|\text{cov}[\Delta_s \ln(1-\tau^p_{it,t-k}), \Delta_s \ln(1-\tau_{it})|\chi_{it}]|$ is large).

Weber (2014) shows that as the lag order increases, the covariance in the numerator of Equation (3) will get arbitrarily close to zero. Mathematically, $\lim_{k\to\infty}|\text{cov}[\Delta_s \ln(1-\tau^p_{it,t-k}), v_{it}|\chi_{it}]| = 0$. But, in contrast to what Weber (2014) suggests, this is not enough to ensure the asymptotic bias in Equation (3) goes to 0. Instead, it is required:

$$\lim_{k\to\infty} \frac{\text{cov}[\Delta_s \ln(1-\tau^p_{it,t-k}), v_{it}|\chi_{it}]}{\text{cov}[\Delta_s \ln(1-\tau^p_{it,t-k}), \Delta_s \ln(1-\tau_{it})|\chi_{it}]} = \frac{\lim_{k\to\infty} \text{cov}[\Delta_s \ln(1-\tau^p_{it,t-k}), v_{it}|\chi_{it}]}{\lim_{k\to\infty} \text{cov}[\Delta_s \ln(1-\tau^p_{it,t-k}), \Delta_s \ln(1-\tau_{it})|\chi_{it}]} = 0. \tag{4}$$

The condition given by Equation (4) states that the covariance between the instruments and the error term must converge to zero *faster* than the covariance between the instruments and instrumented variables. An unfortunate result of the proposed instruments is that as the lag length, k, increases, the explanatory power (relevance) of the instruments will decrease. This is a natural consequence, as using income information further in the past will make it harder to predict actual changes in marginal taxes. As a result, it is not obvious whether Equation (4) will be fulfilled. Consequently, choosing an "appropriate" lag order will yield the lowest asymptotic bias.

In a recent study, Aronsson et al. (2022) confirm the arguments posed by Weber (2014); increasing the lag length of the instrument leads to decreasing bias, although also increasing variance. In general, Aronsson et al. (2022) find that the IV regression estimator works well under "favorable conditions", such as having an analysis period with sufficient variation in marginal taxes and an appropriate structure of the error term in the income process.

Relevancy and validity of the instruments given by Equation (2) will be investigated graphically and by empirical tests. The graphical approach is covered in Section 3.2. To empirically determine relevancy of the instruments, the F-test statistics from the first stage regressions are reported, as is commonly done. The size of these F-test statistics are in all cases above the levels advised to avoid weak instruments (as suggested by e.g. Staiger and Stock (1994)), but the relative sizes of these are still very relevant, as seen by Equation (4). Based on graphical investigation, I choose lag order $k = 2$ in the main specification. In Section 6.4, the validity of the instrument with $k = 2$ is empirically tested using the Difference-in-Sargan over-identification test with instrument lags $k = 2, 3, 4$.

3.1.2 Heterogeneous Growth Rates

Assuming Equation (4) is fulfilled, the issue of mean reversion is solved. If taxpayers at different income levels experience different growth rates in hours of work, this may also introduce bias. Heterogeneous growth rates can arise if e.g. part-time workers or workers paid by the hour, who generally have lower incomes, have higher growth rates in hours of work. If certain income levels are highly consisting of workers close to retirement, these will also have changes in work hours that are different from the overall population. The problem of heterogeneous growth rates is particularly severe when considering tax reforms that are strongly targeting certain income groups. The abolition of the middle-bracket tax in 2010 is a prime example of this. In such cases, the mechanical changes in tax rates described by Equation (2) will be strongly correlated with income level. Other literature control for such heterogeneity by including known background characteristics, but this will be insufficient to absorb the heterogeneous growth rates, if they are not closely related to income or skill levels. The conventional approach has been to include base year income controls. These are endogenous when included directly, as argued by Weber (2014). Instead, the lag order chosen in Section 3.1.1 can be applied to labor income[6] and include this directly. If suitable lags are chosen, the income controls control for permanent income plus an uncorrelated measurement error (Weber (2014)). To allow for non-linearities, I include a ten-piece spline with break-points at the deciles of log labor income lagged twice in my main regressions. The importance of the exact specification of non-linearities is covered in Section 6.2.

3.2 Instruments: Predicted Marginal Net-of-Tax Rates

The analysis period contains 4 tax reforms, which provide the main variation in marginal taxes. This variation is illustrated in Figure B.1 and Figure B.2 by dividing taxpayers into different groups.[7] These groups are based on which tax bracket the taxpayer is located in before and after three years. E.g. "Bottom-middle" refers to a taxpayer going from paying bottom bracket to middle bracket tax. "Mechanical changes" (in navy blue) refer to changes implied by the tax system, calculated as in Equation (2) for $k=0$. "Actual changes" (in maroon red) refer to the actual change occurred in the three year period. In Figure B.1 the share in each grouping is illustrated. It is apparent that is is not possible for taxpayers being mechanically pushed from one bracket to another without a change in bracket thresholds. E.g., for the 2003 reform, the middle-bracket threshold was increased, which mechanically pushed a fair share of taxpayers from the middle to bottom bracket. On the other hand, it can also be seen that a relatively small share of taxpayers are pushed from the top to middle bracket, without the reform specifically targeting the top-bracket threshold. This is due to smaller annual adjustments in the thresholds, which are present for almost all years (see Figure A.1a). Note, the middle-bracket was abolished in 2010 meaning only the combinations of top and bottom brackets for the 2012 reform figure are relevant. Figure B.2 shows the percentage in marginal tax the grouping was subject to in average. Combining both figures Figure B.1 and Figure B.2 shows that the four reforms yielded large changes for a great share of taxpayers. The presence of large heterogeneous tax changes across taxpayers is very apparent. It should also be noted that the mechanical changes explain the actual variation in marginal tax rates very well, when applicable.

A way to graphically examine the validity of the instruments, is to illustrate the change in taxable income for different levels of initial taxable income. The slope of these graphs are of particular interest, as a strong negative slope can be a sign of mean reversion. This is because a negative slope signifies that a low (high) taxable income in a given year is associated with a high (low) difference in taxable income,

[6] In the ETI literature, taxable income is used instead. Whether the controls are based on labor income, taxable income or hours of work in this way does practically no difference for my empirical results.

[7] Note, there are only three figures in Figure B.2 for the four reforms, as the period 2007-2010 captures changes from both the 2007 and 2009 reforms.

exactly as defined by mean reversion. As seen by Figure B.3c, there is a strong, negative correlation between initial taxable income and the log three-year change in log of taxable income. This correlation is particularly severe for initial taxable income lagged 0 years ($k = 0$), and lessens as the lag-order increases. The marginal difference between income lagged 2 and 3 years is small, but even after 3 lags, some of the negative correlation persists. This may either be due to highly persistent serial correlation or due to heterogeneous growth rates, which are difficult to distinguish between. For the main empirical application, mean reversion is only problematic if it comes through hours of work. This is depicted in Figure B.3a, where a negative correlation is again apparent, but it almost entirely disappears after lagging initial labor income twice. For this reason, lag order 2 ($k = 2$) is chosen for the main empirical application in Section 5.

In the ETI literature, it is common to restrict data using an income cut-off, as excluding a certain part of the population can reduce the mean reversion problem (see Auten and Carroll (1999)). A restriction on broad income being above $\approx$ 10,000 USD (1992 level)[8] was popularized by Gruber and Saez (2002). In Figures B.3a-B.3c and the remaining of this paper, a restriction on initial labor income being above 250,000 DKK (2015 level) is imposed. This is a relatively high restriction compared to the one used by Gruber and Saez (2002), but made to ensure conservative estimates. While the restriction only slightly changes main results, it does appear to highly impact the graphical inspection conclusions just made. As seen by comparing Figure B.3d to figures Figure B.3a, imposing a minimum restriction on initial incomes appears to greatly improve the mean reversion problem. The high impact of the restriction suggests significant mean reversion at lower income levels.

4 Data

The data used is based on a combination of different Danish administrative registers. The outcome variable, performed hours of work, is extracted from *Lønstatistikken* (LON). I construct the outcome variable by summing over hours of performed work in all employments in a year per individual to include secondary employments. The data from LON are reported by employers whose contributions to the employees' pensions are based on hours worked. Therefore, there is strong incentive to correctly report them. This register covers around 50 % of main employments in Denmark, but observations can be missing for especially employments in small, private companies. In particular, the register does not cover private firms with less than the equivalent of 10 full-time employees. For reference, see

[8] This corresponds to just below 100,000 DKK (2015 level).

Table A.2 for a description of matches between LON and IDAN, which contains all employments. From the Table A.2 it can be seen that the employments covered in LON are generally in larger firms, both in terms of revenue and number of employees. For an in-depth coverage on *performed* work hours, see Subsection A.1, and for comprehensive quality checks of contracted work hours from LON, see Labanca and Pozzoli (2022). I limit my sample to the period 2002-16, as contract hours are only available from 2002. This should exclude any potential effect from the 1999 reform.

The independent variable of interest, the log change in net of marginal tax rate, has been calculated based on tax simulation programs originally developed by Kleven and Schultz (2014) and subsequently expanded and improved by the Danish Ministry of Finance. These programs are available from 1984-2016. I have adjusted the implementation and collection of relevant data used in the programs to fit the data available to me. Marginal taxes are calculated as $\tau_{it} = [T_t(z_{it} + 100) - T_t(z_{it})]/100$ using income information for the individual gathered from primarily the IND, INDK, and INPI registers. The instrumental variables used are calculated in a similar fashion, but instead use lagged income information (see Section 3). Hence, marginal taxes are calculated as in Kleven and Schultz (2014), while instruments as in Weber (2014). Information used for control variables are gathered from IDAN for workplace controls, BEF and FAIN for various socio-economic controls and municipal controls, FAMILIE for information on children, and UDDA for educational controls.

The primary data restriction is requiring individuals to be of age 25-65, as well as availability of all information needed to calculate marginal taxes. Additionally, it is required that the individual works strictly positive hours and is neither jobless nor outside the labor force as primary occupation as of end November. These conditions are made to exclusively focus on the intensive margin of labor supply, as including the extensive margin would require its own explicit modeling. Descriptive statistics for the sample fulfilling these requirements can be seen from column (1) in Table A.1. All monetary variables have been deflated to 2015 level using the CPI from Statistics Denmark. Subsequently, a condition on base year labor income being higher than 250,000 DKK is imposed, as noted in Section 3.2. Descriptive statistics for the sample after imposing the minimum restriction on taxable income can be seen from column (2) in Table A.1. Finally, restrictions are imposed that all control variables are non-missing for the individual in a given year. This restriction is most influential for lagged income controls, as it requires the individual to be part observable for several periods. Imposing this condition yields the main estimation sample and can be seen from column (3) in Table A.1.

The distribution of performed hours and marginal tax rates faced by individuals in the estimation sample are illustrated in Figure A.3a and Figure A.3b. From Figure A.3a it can be seen that hours of work are relatively evenly distributed, although with a huge spike in mass around 32 hours per week.[9] There is less than expected mass at high hours of work (e.g. $h_{it} \geq 45$), which is a consequence of the unavailability of unpaid overtime. Similarly, the distribution of marginal tax rates are illustrated in Figure A.3b. The mass of this distribution is primarily around the points given by 40, 50, 55, 62 %. These correspond to the marginal tax rate faced by individuals in the bottom, middle and top (post and pre 2009-reform) brackets, as also illustrated in Figure A.2a.

4.1 Graphical Illustration of Variation

In their recent work, Jakobsen and Søgaard (2022) propose a new reduced-form approach that allows for graphical validation of identifying assumptions. By drawing inspiration from their approach, the effect of the 2003 and 2007-09 reforms on performed work hours is illustrated in Figure 1. This should not be interpreted as an alternative to the estimates in Section 5, but more as conceptualized visualization of the effects. While Jakobsen and Søgaard (2022) base their analysis on base-year income, I instead use income lagged twice, as it has a closer resemblance to my main empirical approach. Figure 1 only contains two panels for the 4 reforms in the period 2002-2016. The 2007 and 2009 reforms are pooled in one, and the 2012 reform is left out because of its gradual increases to the top-bracket threshold, which implies a more fuzzy identification of income levels targeted and untargeted by the reform.

Panel (a) of Figure 1 illustrates the percentage change in the predicted net of marginal tax rate using income lagged twice for the 2003 reform (left) and 2007-09 reforms (right). For each reform, the prior three year period is included as a "pre-reform" period, used as comparison. For the 2007-09 reforms I define the "reform" period as 2008-11 to avoid the changes from the 2003 reform in the pre-reform period. This yields a pre-reform period (2005-08) that is relatively stable, which can then be used as a counterfactual to the reform changes. For the 2003 reform, the pre-reform period is much less stable due to the 1999 reform (see Table 1). The dashed red lines indicate which income levels were targeted by the two reforms: The 2003 increased the middle-bracket threshold from $\approx 240,000$ DKK to $\approx 310,000$ (2015 level), while the 2007-09 reforms' primary change is the abolition of the middle-bracket tax, affecting income levels above $\approx 310,000$ DKK (2015 level). There is a spike in net of taxes for relatively high income levels, which is because of small annual changes to the bracket thresholds (see Figure A.1), pushing some individuals from top-bracket to middle-bracket tax.

In panel (b), the three-year change in log hours of performed work is depicted for the two reform

[9] "Hours per week" are calculated as $h_{it} \cdot \frac{7}{365}$ for illustrative purposes. These hours reflect time actually spent at work, meaning the spike in density appears well below 37 hours per week, being the extent of a typical work week.

periods across the income distribution. Using income levels unaffected by reforms (below 240,000 and above 310,000 DKK for the 2003 reform and income below 310,000 DKK for the 2007-09 reforms) as control group, the differences in hours of work between the reform period and pre-reform periods can be interpreted as the reform effect under a parallel trends assumption. Panel (b) for the 2007-09 reforms suggest that a tax cut's potential positive effect on hours of work need not come from an increase in hours of work, but can instead be caused by a smaller decrease than in absence of the reform. This could e.g. be explained by the tax cut incentivizing older workers close to retirement to keep their current work hours fixed for longer periods, instead of decreasing them.

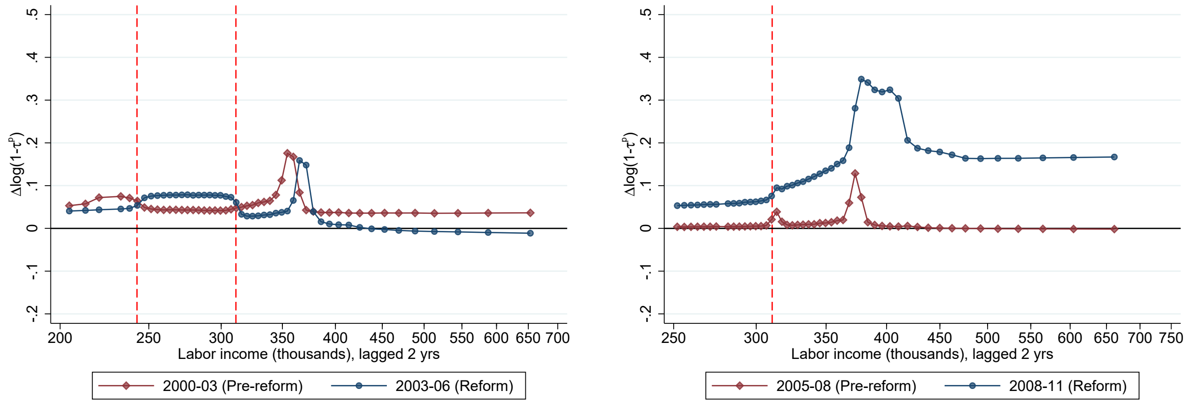
Panel (c) illustrates the relative change in hours of performed work between the pre-reform and reform periods. For the 2007-09 reforms, the "pre-treatment" (income below 310,000) treatment effect is statistically insignificant at the 5 % level, but also appears unstable, which challenges the parallel trends assumption. If one has confidence in the assumption being fulfilled, the trend differentials to the right of the dashed red line can be converted to elasticities by scaling them with the appropriate change in marginal taxes from panel (a). The result would be elasticities close to 0 for individuals with income just above the cut-off, with much higher elasticities for higher incomes. For the 2003 reform, the "treatment" area is much narrower, and while there appears to be small increases in hours of performed work, these are all insignificant at the 5 % level. In any case, a causal interpretation would here be hard to validate due to the unstable "control group", as seen in panel (a).

Figure 1 provides evidence for the existence of potential causal effects of tax rate changes on labor supply. The figure also suggests responsiveness in hours of work of high earners for the 2007-09 reforms, for which there has been very scarce documentation in the literature (Moffitt and Wilhelm, 2000; Saez et al., 2012). Additionally, unstable tax systems in "pre-reform" periods and a questionable parallel trends assumption highlights the need for the regression based approach used in Section 5.

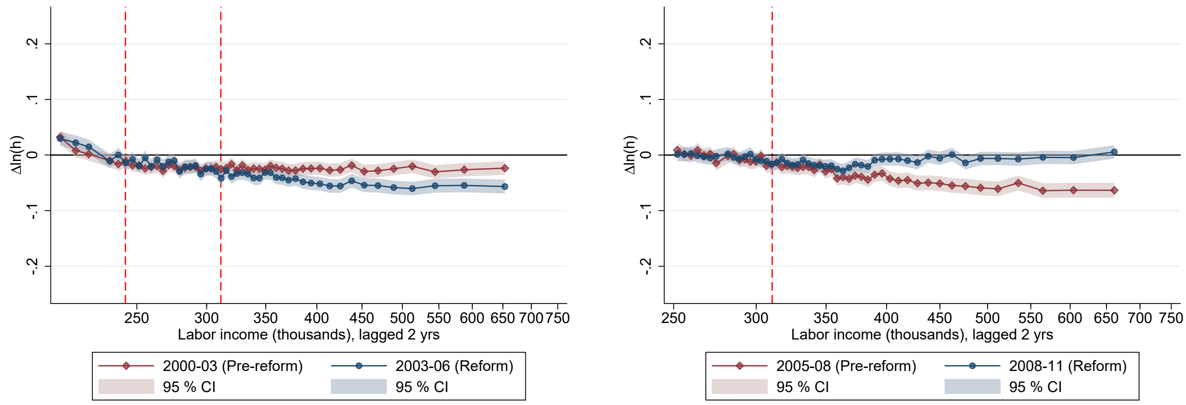
Figure 1: Graphical Illustration of Labor Responses

2003 Reform | 2007-09 Reforms

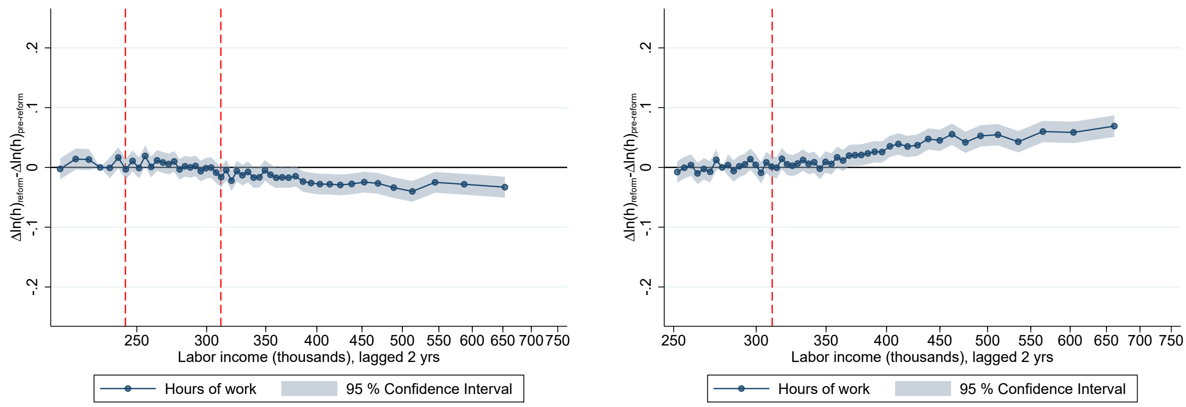
(a) Changes in Predicted Net-of-Tax Rates | (a) Changes in Predicted Net-of-Tax Rates



(b) Trends in Work Hours Across Income | (b) Trends in Work Hours Across Income



(c) Changes in Trend Differentials | (c) Changes in Trend Differentials



Panel (a) shows the change in predicted log net-of-tax rates (Equation (2) for $k = 2$), in panel (b) the change in *performed* hours of work, and in panel (c) the changes in trend differentials in hours of performed work, comparing pre-reform to post-reform periods for the 2003 reform (left) and 2008 reform (right). The red dashed line indicate income levels the reforms target. The 2003 reform targets taxpayers in the middle-bracket (income $\approx$ between 240.000 and 310.000 DKK (2015 level)), while the 2008 reform abolished the middle-bracket tax, affecting taxpayers with income above $\approx$ 310,000 DKK (2015 level).

5 Results

The empirical estimates of Equation (1) in three-year changes can be seen from Table 2, where the preferred specification estimates are displayed in column (4). Applying several different changes to the main specification is covered in Section 6 and first stage regressions are provided in Table B.1.

The four columns of Table 2 all include controls for year fixed effects, municipality fixed effects, and socio-economic controls. The last row contains the F-statistics from testing excluded instruments. The specification in column (1) yields a negative and statistically significant ELS, estimated to be -0.0214. Column (2) includes a 10-piece spline of log base year labor income and the estimate turns positive, still statistically significant at the 1 % level. This resembles the results of Kleven and Schultz (2014). Including controls for base year income has been the standard way to control for both mean reversion and heterogeneous growth rates simultaneously in the ETI literature and column (2) illustrates the importance of doing so. However, assuming that the transitory income component is serially correlated, this control for base year income is invalid. In column (3) instruments instead are based on lagged income, as proposed in Section 3, where the elasticity is still positive, but statistically insignificant and much smaller in size. Moving from column (3) to (4), a 10-piece spline of base year log labor income lagged twice is included, this time only to control for heterogeneous growth rates. By lagging income two times, the serial correlation should be "played out", and hence the controls for heterogeneous growth rates used in column (4) should not suffer from the same problem as in column (2). Doing so, the estimated ELS is positive, of considerable size, and statistically significant at the 1 %-level. The estimate of 0.0814 implies that a decrease in the marginal tax rate of 10 % leads to an increase in labor supply of $\approx$ 0.8 % at the intensive margin. The size of the estimated elasticity is relatively in line with with previous labor supply studies (0.15 by MaCurdy (1981), 0.13-0.37 by Blundell et al. (1998), 0.13 by Ziliak and Kniesner (1999), 0.14 by Bingley and Lanot (2002), approx. 0 for Denmark by Bargain et al. (2014)), albeit at the lower end.

Table 2: Labor Supply Elasticity - Main Regression Results

	(1) $\Delta \ln h$	(2) $\Delta \ln h$	(3) $\Delta \ln h$	(4) $\Delta \ln h$
$\Delta \ln(1-\tau)$	-0.0214***	0.0241***	0.0038	0.0814***
	(0.0028)	(0.0032)	(0.0078)	(0.0091)
First stage F-statistic	1442042	1327015	145300	108445
Time fixed effects	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes
Socio-economic controls	Yes	Yes	Yes	Yes
Instrument lags	0	0	2	2
Income controls	None	10-piece spline	None	10-piece spline, lag 2
Observations	10,755,733	10,755,733	10,755,733	10,755,733

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

Labor supply elasticities estimated by 2SLS regressions. The dependent variable in all specifications is the three-year log change in *performed* hours of work. The independent variable of interest is in all cases the three-year log change in the marginal tax on additional labor income. This has been instrumented by the mechanical changes, calculated by Equation (2), where k is given by the row "Instrument lags". The row "Income controls" denotes which income controls are included. "10-piece spline" refers to a 10-piece spline with breakpoints at deciles of log labor income. All specifications include time, individual fixed effects, socio-economic controls, and municipality fixed effects. The socio-economic controls are as listed in Section 3.

5.1 Decomposing the Variation

The variation in hours of work comes from two channels: within-job variation and across-job variation. Table 3 and 4 explore these two channels in more detail.

In Table 3, the main regression specification (column (4) of Table 2) is re-estimated, imposing different job-switch conditions. Firstly, column (1) of Table 3 repeats the results of column (4), Table 2. I.e., it is unconditional on job switches and serves as the benchmark. Column (2) in table Table 3 conditions on the individual not having a new job in period *t+3*, compared to period *t*. This column yields the within-job variation in hours of work. The estimated elasticity is smaller compared to column (1), but still sizable. This suggests that there is a lower variation in within-job hours of performed work, compared to the overall variation. In column (3), the elasticity is estimated conditional on the individual changing main employment from period *t* to *t+3*. Comparing column (3) to (2), the estimated elasticity is more than twice as large and the two are statistically different at the 5 % level. Finally, column (4) conditions on the individual switching or taking up new secondary employment from period *t* to *t+3*. The resulting elasticity is also higher than in columns (1) and (2), albeit now statistically significant only at the 5 %-level. Note, the number of observations in columns (2)-(4) do not add up to the number of

Table 3: Results Decomposed - Imposing Job Switch Conditions

	(1) $\Delta \ln h$ No condition	(2) $\Delta \ln h$ No new job	(3) $\Delta \ln h$ New main job	(4) $\Delta \ln h$ New secondary job
$\Delta \ln(1-\tau)$	0.0814***	0.0566***	0.1279***	0.0988**
	(0.0091)	(0.0083)	(0.0287)	(0.0496)
First stage F-statistic	108445	72466	17898	7087
Time fixed effects	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes
Socio-economic controls	Yes	Yes	Yes	Yes
Observations	10,755,733	6,030,719	3,717,794	2,000,044

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

Labor supply elasticities estimated as in column (4) of Table 2. The dependent variable in all specifications is the three-year log change in hours of work. Columns (2)-(4) each impose conditions on job changes for the three-year period. Column (1) does not impose any conditions, is equivalent to column (4) of Table 2, and serves as the benchmark. Column (2) imposes that the individual does not have a new job at time *t+3*, compared to time *t*, while column (3) imposes the individual has a new *main* job, and column (4) imposes a new *secondary* job. These conditions are in most cases mutually exclusive, but not necessarily so.

observations in column (1), as individuals may change both main and secondary employment, resulting in them appearing in both column (3) and (4) simultaneously. It should also be noted that imposing the job-switch conditions can be interpreted as conditioning on post-treatment controls, which may bias results if tax reforms influence job switching probability, which is likely indeed the case, as documented by e.g. Gentry and Hubbard (2004); Kreiner et al. (2015). This potential bias is ignored here. The results summarized in Table 3 suggests that a large part of the elasticity estimated in the main regression results can be attributed to job changes. This has important policy implications, because it suggests that the possibility and ease of job-switches is an important determinant of the existence and magnitude of real behavioral responses to tax rate changes.

The results presented in Table 4 further investigate what type of hours the variation in performed hours comes from. Importantly, note that actual changes in the outcome hours are estimated instead of log changes. This is done by using a semi-log specification, changing the interpretation of the results. This is to not overemphasize large relative changes in absence and overtime hours, which are at a much lower mean level than performed work hours. Using the log-log specification as stated in Equation (1) would also requires a strictly positive number of hours in both year *t* and *t+3*, which is very restrictive for absence and overtime. Changing the functional form of estimation may introduce some specification bias, but is here assumed to be negligible. With this in mind, column (1) of Table 4 estimates a 1 % reduction in the marginal tax rate leads to an increase in performed hours work by 1.85 hours.[10]

[10] With a mean of ≈ 29.04 performed work hours in the given sample, this yields elasticity relatively close to that in Table 2.

Table 4: Results Decomposed - Type of Work Hours

	(1)	(2)	(3)	(4)
	Δh	Δh^c	Δh^a	Δh^o
	Performed hours	Contract hours	Absence hours	Overtime hours
$\Delta \ln(1-\tau)$	1.8480***	1.4114***	-1.1107***	0.0842***
	(0.1477)	(0.1695)	(0.0706)	(0.0312)
First stage F-statistic	108431	108431	108431	108431
Time fixed effects	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes
Socio-economic controls	Yes	Yes	Yes	Yes
Observations	10,753,860	10,753,860	10,753,860	10,753,860

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

Labor supply elasticities estimated as in column (4) of Table 2. Note, here a condition is imposed to ensure availability of all outcome variables. The dependent variable is the three-year change in different measures of work hours. In column (1) the dependent variable is based on hours of performed work, column (2) is contract hours, column (3) is hours of paid absence, and column (4) is hours of overtime work.

Performed hours is contract hours (column (2)) minus paid absence hours (column (3)) plus overtime hours (column (4)) minus vacation and holiday hours. From columns (2)-(4) of Table 4, it is seen that a large fraction of the variation in performed hours comes from variation in contract work hours, suggesting that individuals obtain more contracted work hours. This can be done by going from part-time to full-time employment or switching to a workplace that offers more work hours. However, there is also a large reduction in paid absence (column (3)).[11] A smaller part of the increase in performed hours also seems to come from paid overtime, as seen from column (4). The small size of this effect is most likely small due to workplaces rarely utilizing paid overtime. The decrease in paid absence may be caused by workers making less use of paid sick leave, having their "sixth vacation week" paid out as wage increase instead of spending it, taking less maternity leave or similar. The reduction in paid absence may also be caused by intra-household substitution of paid absence, where the person subject to a tax cut makes less use of paid absence, e.g. due to child sickness, and instead their spouse spends this absence.

Decomposing the variation in hours of work suggests the importance of labor market fluidity and flexibility for labor supply responsiveness, by making job changes and adjustments to paid absence and overtime easier. This can potentially help explain cross-country differences, with Denmark having a both fluid (see e.g. Engbom (2022)) and flexible labor market. These findings further emphasize the importance of contextual factors for elasticity estimates, similarly to the findings of Neisser (2021). Considering that the rate of labor market turnover and employment protection in Denmark is comparable

[11]For absence hours, the caveat of imputed values noted in Subsection A.1 should be in borne in mind.

to those of the United Sates (Kreiner and Svarer, 2022), the points made in this paper may very well carry over to the United States.

5.2 Relating Labor Supply, Labor Income, and Taxable Income Elasticities

Given the close connection to the ETI literature, the main specification used in column (4) of Table 2 can also be used to estimate the elasticity of labor income (ELI) and taxable income. This is done in Table 5. It should be noted the methodological decisions have been made with estimating the elasticity of labor supply as primary objective. This is particularly relevant for choosing the appropriate lag order, k, in Section 3.1.1. Table 5 uses the same conditions as in the main results, but with the additional condition of lagged taxable income being available, which is used in column (3). The ELS, ELI, and ETI are estimated in column (1) to (3), respectively. Income controls in column (3) are based on log taxable income, as opposed to log labor income in columns (1) and (2). Comparing column (1) to (2), it can be seen that the ELS is actually *larger* than ELI. This is very likely because *performed* hours of work does not include all paid work hours. In particular, a decrease in paid absence will increase *performed* work hours, but not directly affect labor income. However, it should be noted that general equilibrium effects such as a decreased wage level due to higher aggregate labor supply in the economy could also contribute to the explanation (as found by e.g. Blomquist and Selin (2010)), although ignored in this reduced form framework. Comparing column (1) to (2) provides evidence that the primary driver of ELI is from individuals adjusting their work hours, and not from individuals switching to better paying jobs, more actively negotiating for wage increases and promotions, or similar. The ETI in column (3) includes all behavioral responses and is the highest of the three elasticities. ETI being larger than ELI may be driven by income shifting from certain pension contributions (which are deducted in taxable income, see Section 2) to personal income as documented for the 2008 reform by Jakobsen and Søgaard (2022). Table 5 provides evidence that the majority of the total response can be attributed to real responses, as opposed to avoidance and timing responses. ELS being the primary driver of behavioral responses contributes to the findings of Kleven and Schultz (2014), who argue that low taxable income elasticities in Denmark is due to small opportunities for avoidance and evasion.

Table 5: Labor Supply, Labor Income, and Taxable Income Elasticities

	(1) $\Delta \ln(\mathrm{h})$	(2) $\Delta \ln(\text{labor income})$	(3) $\Delta \ln(\text{taxable income})$
$\Delta \ln(1-\tau)$	0.0817***	0.0579***	0.0952***
	(0.0091)	(0.0037)	(0.0042)
First stage F-statistic	112346	112346	128227
Time fixed effects	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes
Socio-economic controls	Yes	Yes	Yes
Observations	10,683,225	10,683,225	10,683,225

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

Elasticities estimated as in column (4) of Table 2. The dependent variable in column (1) is the three-year log change in *performed* hours of work, in column (2) the three-year log change in labor income, and in column (3) the three-year change in taxable income. Note, the income splines in column (3) are based on log taxable income, opposed to log labor income as in columns (1) and (2).

6 Robustness Checks

6.1 Lag Lengths and Difference Sizes

In the results so far, I have considered a difference period of 3 years with instruments and income controls lagged 2 years. To assess the sensitivity to this choice, Table B.2 and B.3 repeat the main regression estimates, but with varying difference lengths and lags.

Table B.2 shows that increasing the difference length also increases the estimated elasticity up until a certain point, after which the elasticity decreases again. This can be explained by an increased difference length being able to account for sluggishness in behavioral responses and capturing long-term effects, while increasing the difference length too much leads to including noise and higher data requirements. Going from column (1) to (2), the elasticity disappears entirely, which is a peculiar result. This could be caused by the two year difference not being long enough to overcome behavioral frictions while still including overlapping intervals. This is especially a problem for tax reforms with phase-in-periods or closely placed tax reforms, which the 2007 and 2009 reforms are prime examples of. Moving across columns, the number of observations decreases sharply, as calculating differences over longer periods is increasingly data demanding. Here, the problem of using overlapping differences should also be noted; the estimates are a combination of short-run, medium-run, and longer-run responses. E.g., in column (2), the 2002-04 difference will estimate the the short-run effect of the 2003 reform, while the 2003-05 difference will capture longer-run responses, and the 2004-06 difference will capture responses if these are sluggish, even if no tax rate change is present. In addition, it should be noted that not even six-year differences may be able to capture all long-run responses, such as through choice of education or upgrading labor skills and human capital.

From Table B.3 it is clear that increasing the lag order increases the size of the estimated elasticity monotonically. This is similar to the results of Weber (2014), whom argue it is caused by higher exogeneity of the instruments with increasing lag orders. However, as seen in Section 3.2, increasing the lag order leads to very small visual changes of mean reversion. Hence, the difference is more likely to be caused by two things. Firstly, the weaker first stage when increasing the lag order might imply the denominator of Equation (3) converges faster to zero than the numerator, actually *increasing* bias. Secondly, further lagging the instruments may shift weights given to compliers, always takers and never takers in the LATE-interpretation of the IV estimator. The compliant population may also change to individuals who are more responsive to tax reforms, as the lag order increases. The first point emphasizes the importance of the graphical inspection made in Section 3.2 and that instruments should only be lagged if the degree of mean reversion is sufficiently decreased.

Overall, while the main elasticity estimate is dependent on the choice of difference size and instrument lag, it is not overly sensitive. The results in Table B.3 suggests careful consideration should be made when choosing the appropriate lag-order.

6.2 Alternative Income Controls

The ETI literature has previously suffered from sensitivity to non-linearity in income controls (see e.g. Kopczuk (2005)). To investigate the robustness of choosing a 10-piece spline in the main estimates, I now re-estimate these with differing functional forms. This is done in table Table B.4, where column (3) repeats column (4) of Table 2. The resulting estimates show that controlling for non-linearities is important, as seen by going from column (1) to (2) where a 5-piece spline of log labor income is included instead of log labor income. But once taking these non-linearities into account, the results are very robust to increasing the degree of non-linearities. This is seen by the small change in elasticity estimate from increasing the spline pieces to 10, 20 and by including controls for deviation between log labor income between year $t-2$ and t. The importance of controlling for non-linearities is well explained by the tax-reforms targeting specific income groups as well as heterogeneous growth rates. While it is important to control for non-linearities, the exact specification used is much less so.

6.3 Sensitivity to Inclusion of Control Variables

In Table B.5, the sensitivity to inclusion of control variables is examined. Going from column (1) to (4), an increasing degree of control variables are included, where column (4) is the main specification. The largest difference in elasticity estimate comes from including human capital controls, going from column (1) to (2). Although not directly seen from the table, the most important factor to control for is experience, which is in line with the findings of Keane (2011, 2021). In column (5) of Table B.5, the 10 spline-segments are interacted with year dummies to test the sensitivity to changes in heterogeneous growth rates over time. Comparing column (4) with (5), there is a slight decrease in the estimated elasticity, but they are not significantly different. Lastly in column (6), I use a "residualized" instrument, as suggested by Kumar and Liang (2020), as they argue it leads to lower bias and higher efficiency. The residualized instrument is derived by regressing the marginal tax instrument on a ten-piece spline on initial labor income lagged twice and using the residuals as instrumental variable instead of including income controls directly in the regression. As seen by comparing column (4) to (6), this yields practically no difference.

6.4 Empirically Testing the Validity of Instruments

The Difference-in-Sargan test (also called "GMM distance" or C-test) is used in Table B.6 to empirically test the validity of the instruments. The Difference-in-Sargan test is an over-identification test that allows testing the exogeneity of one or more instruments. The test is defined as the difference in the Sargan-Hansen (J) statistic when calculated using the restricted model (including only the instruments valid under both the null and alternative hypotheses) and the unrestricted model (including all instruments, also those whose validity is suspect). Under the null, the suspect instruments are valid and the resulting C statistic is χ^2-distributed with $m-\kappa$ degrees of freedom, m being the total number of instruments, κ being the number of suspect instruments (Hayashi, 2000). I will use $m = 3$ and $\kappa = 1$, testing the validity of $\Delta_s \ln(1-\tau^p_{it,t-k})$ by using $\Delta_s \ln(1-\tau^p_{it,t-k-1})$ and $\Delta_s \ln(1-\tau^p_{it,t-k-2})$. In the main specification, $k = 2$. When $m-1 = \kappa$ the Difference-in-Sargan test is equivalent to the Sargan test (also called "J" or Hausman test) (Newey, 1985). The advantage of the Difference-in-Sargan test over the Sargan test is that it has been found to have more power (Arellano and Bond, 1991). Employing this strategy to the main estimation results (Table 2), is seen from Table B.6. In columns (1) and (2), the exogeneity of the instrument lagged 0 times is tested using instruments lagged 1 and 2 times, while in columns (3) and (4) the exogeneity of the instrument lagged 2 times is tested using the instrument lagged 3 and 4 times. The P-value from the Difference-in-Sargan test is presented at the bottom of each column. Interestingly, exogeneity is not rejected in column (1) without income controls, while it is in column (2). The reverse holds true comparing column (3) with (4). This pattern may emerge due to the inclusion of non-valid income controls used in column (2) and the difficulty to distinguish between highly persistent serial correlation in the income process and heterogeneous growth rates, as argued in Section 3.2. Importantly, exogeneity of the instrument used in the main specification (column (4)) is not rejected for P-values below 30 %. Choosing only the instrument lagged twice in Table 2 is essentially following the advice of Angrist and Pischke (2008): Picking the best single instrument and reporting the just-identified estimates.

7 Conclusion

Given it is possible to design the tax system to reduce avoidance opportunities and increase tax enforcement, the type of behavioral responses to taxation is crucial (Saez et al., 2012). If avoidance and timing responses can successfully be minimized, real responses, in the form of labor supply, is the bedrock of behavioral responses to taxation.

In this paper, I estimate the elasticity of labor supply (ELS) to be 0.08 using a series of Danish tax reforms over the period 2002-16 and by applying conventional ETI methodology to large administrative data on performed hours of work. Keeping a close connection to the ETI methodology allows me to additionally estimate the ETI, which I estimate to be 0.1, and thereby show that a large fraction of total responses can be attributed to responses in hours of work.

The ELS is primarily driven by individuals switching main and secondary employment. For individuals not switching jobs over the three-year period analyzed, the elasticity is more modest, but still significant. The effect on performed work hours comes from both an increase in contracted hours and a decrease in paid absence. An increase in paid overtime also contributes to the total effect but is smaller, likely as the use of paid overtime is not widespread in Danish firms.

I argue that labor supply responses are the main driver of ETI in Denmark, firstly, because avoidance and timing responses are low for wage earners who are subject to third-party reporting in Denmark (Kleven et al., 2011; Kleven and Schultz, 2014). Additionally, that ELS is primarily driven by individuals switching main and secondary employment suggests that the possibility and ease of job-switches is an important determinant of the existence and magnitude of real behavioral responses to tax rate changes. That job-switches play a large role for ELS suggests that this should be higher for countries with higher labor market fluidity such as Denmark, the United States and the United Kingdom (see e.g. Engbom (2022)) and can potentially explain cross-country variation. Assuming the labor market conditions play an important role for the ELS, many of the conclusions made in this paper can be carried over to the United States, as it exhibits roughly the same labor market turnover and employment protection as Denmark (Kreiner and Svarer, 2022). The marginal tax rate's effect on paid absence and paid overtime work further suggests that job flexibility may be important for labor supply responses. These results further emphasize the importance of contextual factors for elasticity estimates, similarly to the findings of Neisser (2021). A similar decomposition of the total responses would be very beneficial for other countries to further explore cross-country differences and crucial contextual factors, although near impossible because of the high data requirements.

Future work may follow several avenues. An obvious and important caveat of this study is the unavailability of unpaid overtime work. There is no immediate solution to this and other authors have struggled with finding an appropriate measure of work time for decades. I show that a sizable and statistically significant elasticity can be found using administrative register data. This is promising for future studies, which may closely follow the improvements in register-based measures of work time. The approach in this paper can be extended by estimating compensated elasticities similarly to Gruber and Saez (2002) and Kleven and Schultz (2014). Future work may examine the extensive margin or focus on estimating other components of the ETI.

Bibliography

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics*. Princeton University Press.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297.

Aronsson, T., Jenderny, K., and Lanot, G. (2022). The quality of the estimators of the ETI. *Journal of Public Economics*, 212:104679.

Auten, G. and Carroll, R. (1999). The effect of income taxes on household income. *Review of Economics and Statistics*, 81(4):681–693.

Bækgaard, H. (2014). The differences-in-differences approach with overlapping differences - experimental verification of estimation bias. *DREAM Publications*.

Bargain, O., Orsini, K., and Peichl, A. (2014). Comparing labor supply elasticities in Europe and the United States - new results. *Journal of Human Resources*, 49(3):723–838.

Bingley, P. and Lanot, G. (2002). The incidence of income tax on wages and labour supply. *Journal of Public Economics*, 83(2):173–194.

Blomquist, S. and Selin, H. (2010). Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates. *Journal of Public Economics*, 94(11-12):878–889.

Blundell, R., Duncan, A., and Meghir, C. (1998). Estimating labor supply responses using tax reforms. *Econometrica*, 66(4):827–861.

Blundell, R. and Macurdy, T. (1999). Labor supply: A review of alternative approaches. *Handbook of Labor Economics*, 3(1):1559–1695.

Burns, S. K. and Ziliak, J. P. (2017). Identifying the elasticity of taxable income. *The Economic Journal*, 127(600):297–329.

Chetty, R. (2009). Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance. *American Economic Journal: Economic Policy*, 1(2):31–52.

Doerrenberg, P., Peichl, A., and Siegloch, S. (2017). The elasticity of taxable income in the presence of deduction possibilities. *Journal of Public Economics*, 151:41–55.

Engbom, N. (2022). Labor market fluidity and human capital accumulation. *NBER (working paper)*.

Fack, G. and Landais, C. (2016). The effect of tax enforcement on tax elasticities: Evidence from charitable contributions in France. *Journal of Public Economics*, 133:23–40.

Feldstein, M. (1995). The effect of marginal tax rates on taxable income: A panel study of the 1986 Tax Reform Act. *Journal of Political Economy*, 103(3):551–572.

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

Frederiksen, A., Graversen, E. K., and Smith, N. (2001). Overtime work, dual job holding and taxation. *SSRN (working paper)*.

Gelber, A. M. and Mitchell, J. W. (2012). Taxes and time allocation: Evidence from single women and men. *The Review of Economic Studies*, 79(3):863–897.

Gentry, W. M. and Hubbard, R. G. (2004). The effects of progressive income taxation on job turnover. *Journal of Public Economics*, 88(11):2301–2322.

Goolsbee, A. (2000). What happens when you tax the rich? Evidence from executive compensation. *Journal of Political Economy*, 108(2):352–378.

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

Harju, J. and Matikka, T. (2016). The elasticity of taxable income and income-shifting: What is “real” and what is not? *International Tax and Public Finance*, 23(4):640–669.

Hayashi, F. (2000). *Econometrics*. Princeton University Press.

Holmlund, B. and Söderström, M. (2008). Estimating dynamic income responses to tax reforms: Swedish evidence. Technical report, Working Paper.

Jakobsen, K. M. and Søgaard, J. E. (2022). Identifying behavioral responses to tax reforms: New insights and a new approach. *Journal of Public Economics*, 212:104691.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature*, 49(4):961–1075.

Keane, M. P. (2021). Recent research on labor supply: Implications for tax and transfer policy. *Labour Economics*, 77:102026.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., and Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3):651–692.

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

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

Kreiner, C. T., Leth-Petersen, S., and Skov, P. E. (2014). Year-end tax planning of top management: Evidence from high-frequency payroll data. *American Economic Review*, 104(5):154–58.

Kreiner, C. T., Leth-Petersen, S., and Skov, P. E. (2016). Tax reforms and intertemporal shifting of wage income: Evidence from Danish monthly payroll records. *American Economic Journal: Economic Policy*, 8(3):233–257.

Kreiner, C. T., Munch, J. R., and Whitta-Jacobsen, H. J. (2015). Taxation and the long run allocation of labor: Theory and Danish evidence. *Journal of Public Economics*, 127:74–86.

Kreiner, C. T. and Svarer, M. (2022). Danish flexicurity: Rights and duties. *Forthcoming: Journal of Economic Perspectives*.

Kumar, A. and Liang, C.-Y. (2020). Estimating taxable income responses with elasticity heterogeneity. *Journal of Public Economics*, 188:104209.

Labanca, C. and Pozzoli, D. (2022). Constraints on hours within the firm. *Journal of Labor Economics*, 40(2):473–503.

MaCurdy, T. E. (1981). An empirical model of labor supply in a life-cycle setting. *Journal of political Economy*, 89(6):1059–1085.

Meyer, B. D. and Rosenbaum, D. T. (2001). Welfare, the earned income tax credit, and the labor supply of single mothers. *The Quarterly Journal of Economics*, 116(3):1063–1114.

Moffitt, R. A. and Wilhelm, M. (2000). *Chapter 7: Taxation and the Labor Supply: Decisions of the Affluent, In: Slemrod J. (Ed.) Does Atlas Shrug? The Economic Consequences of Taxing the Rich.* Cambridge University Press, Cambridge.

Neisser, C. (2021). The elasticity of taxable income: A meta-regression analysis. *The Economic Journal*, 131(640):3365–3391.

Newey, W. K. (1985). Generalized method of moments specification testing. *Journal of Econometrics*, 29(3):229–256.

Saez, E., Slemrod, J., and Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50(1):3–50.

Slemrod, J. (1990). *Do Taxes Matter?: The Impact of the Tax Reform Act of 1986*. The MIT Press.

Slemrod, J. (1995). Income creation or income shifting? Behavioral responses to the Tax Reform Act of 1986. *The American Economic Review*, 85(2):175–180.

Slemrod, J. and Yitzhaki, S. (2002). *Chapter 22: Tax Avoidance, Evasion, and Administration, In: Auerbach, A. J and Feldstein, M. (Eds.), Handbook of Public Economics*, volume 3. Elsevier.

Staiger, D. O. and Stock, J. H. (1994). Instrumental variables regression with weak instruments. Technical report, Working Paper.

Vattø, T. E. (2020). Estimating the elasticity of taxable income when earnings responses are sluggish. Technical report, Working Paper.

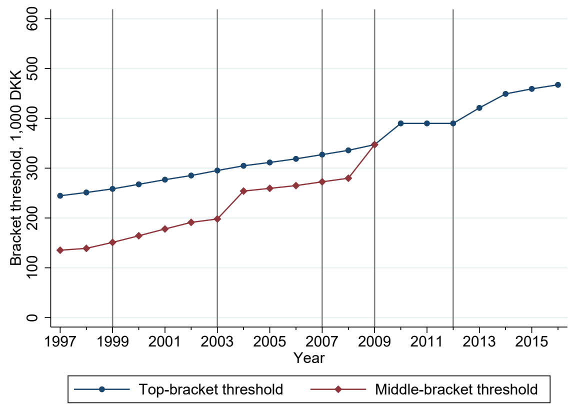
Weber, C. E. (2014). Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. *Journal of Public Economics*, 117:90–103.

Ziliak, J. P. and Kniesner, T. J. (1999). Estimating life cycle labor supply tax effects. *Journal of Political Economy*, 107(2):326–359.

A Descriptives

Figure A.1: Bracket Thresholds and Tax Rates

(a) Bracket Thresholds Over Time



(b) Bracket Tax Rates Over Time

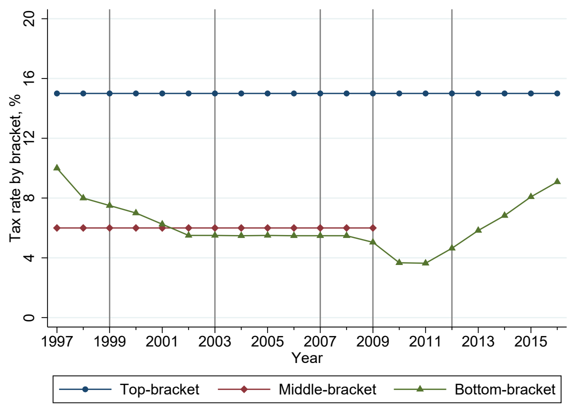
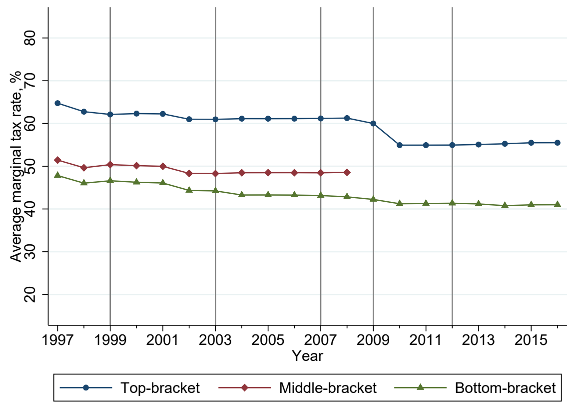


Figure A.1a shows the bracket thresholds in nominal DKK over time. Figure A.1b shows the tax rate in each bracket over time. The vertical grey lines denote the tax reforms as seen in Table 1. The large increase in the bottom-bracket tax rate starting from 2011 was accompanied by an equally large decrease in health contributions, leaving their sum unchanged.

Figure A.2: Marginal Tax Rates and Percent of Population By Tax Brackets

(a) Average Marginal Tax by Tax Brackets



(b) Percent of Population in Each Tax Bracket

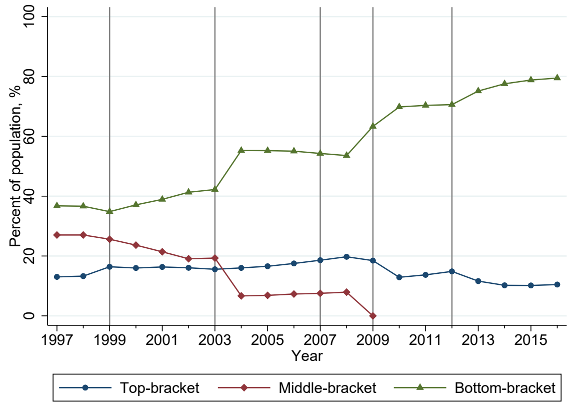


Figure A.2a shows the average marginal tax on additional labor income for tax-payers in the bottom, middle or top tax bracket over time. Figure A.2b shows the percent of the population in each bracket over time. Taxpayers are defined to be in the bottom bracket if they pay bottom bracket tax and neither middle nor top bracket tax. Similarly, taxpayers are in the middle bracket if they pay bottom and middle bracket tax, but not top bracket tax. Taxpayers are defined to be in the top-bracket if they pay top bracket tax. The shares in Figure A.2b do not add up to 100 %, as some taxpayers only pay labor market contributions. The vertical grey lines denote the tax reforms as seen in Table 1.

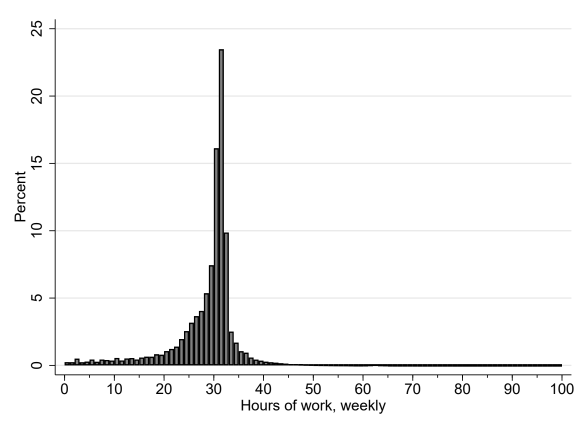
Table A.1: Descriptive Statistics

	(1) All		(2) Labor income $\geq$ 250,000		(3) Estimation sample	
	Mean	Std.	Mean	Std.	Mean	Std.
Performed hours of work	28.18	7.23	29.02	6.35	29.04	6.33
Marginal tax, τ (percent)	49.61	8.68	50.26	8.44	50.27	8.43
Women (fraction)	0.52	0.50	0.51	0.50	0.51	0.50
Married (fraction)	0.62	0.48	0.63	0.48	0.63	0.48
Age (years)	43.36	9.35	43.82	9.21	43.88	9.20
Experience (years)	19.62	9.90	20.23	9.74	20.37	9.66
Has child 0-2 years (fraction)	0.12	0.32	0.11	0.31	0.11	0.31
Has child 3-6 years (fraction)	0.16	0.36	0.15	0.36	0.15	0.36
Has child 7-9 years (fraction)	0.13	0.34	0.13	0.34	0.13	0.34
Has child 10-14 years (fraction)	0.20	0.40	0.20	0.40	0.20	0.40
Has child 15-17 years (fraction)	0.13	0.34	0.14	0.34	0.14	0.34
Education $<$ high school (fraction)	0.13	0.34	0.13	0.34	0.13	0.34
High school education (fraction)	0.43	0.49	0.42	0.49	0.43	0.49
College graduate (fraction)	0.43	0.50	0.44	0.50	0.44	0.50
Labor income, thousands DKK	407.08	180.00	425.40	173.79	425.59	173.76
Taxable income, thousands DKK	3137.85	1589.76	3256.37	1583.70	3257.36	1582.04
Bottom tax bracket (fraction)	0.47	0.50	0.44	0.50	0.44	0.50
Top tax bracket (fraction)	0.38	0.49	0.41	0.49	0.41	0.49
Observations	11,757,144		10,874,515		10,755,733	

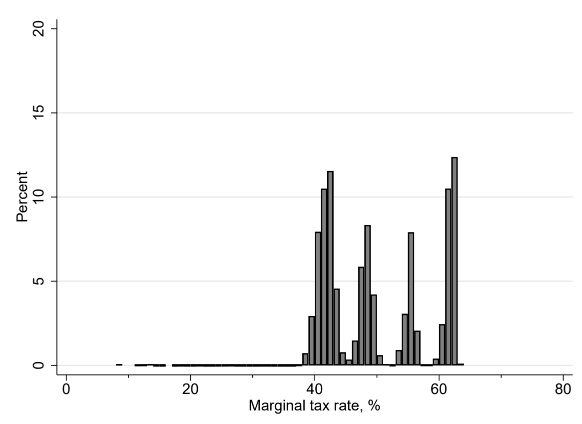
Descriptive statistics for main variables of interest for the analysis period, 2002-16. Annual performed hours of work have been converted approximately weekly by multiplying them with $\frac{7}{365}$. Column (1) contains descriptives for the initial sample, only imposing that individuals are not unemployed or outside the labor force, have strictly positive work hours at time t & t+3, non-missing $\Delta \ln(1-\tau)$, and more than one observation in the sample period. Column (2) additionally conditions labor income being greater than 250,000 DKK. Finally, in column (3) it is imposed that all variables included in the main regressions (column (4) of Table 2) are non-missing. All monetary variables are measured in 2015 DKK.

Figure A.3: Histograms for Performed Hours of Work and Marginal Taxes

(a) Histogram of Performed Hours of Work



(b) Histogram of Marginal Taxes



Histograms showing the distribution of performed hours of work (left) and marginal taxes (right). Each bin has width 1 and is conditional on containing > 5 observations. Otherwise it has been dropped. The performed hours of work in Figure A.3a are made weekly by multiplying the annual hours with $\frac{7}{365}$.

Table A.2: Matches between *IDAN* and *LON*

(1) Year	(2) % matched, IDAN-LON	(3) Avg. employees, IDAN	(4) Avg. employees, LON	(5) Avg. revenue (M DKK), IDAN	(6) Avg. revenue (M DKK), LON
1997	46.53	.	.	15.32	48.85
1998	47.21	.	.	15.46	49.66
1999	47.25	13	37	14.70	46.59
2000	54.17	16	49	16.48	50.81
2001	53.99	15	44	15.01	45.07
2002	54.02	16	47	16.08	47.49
2003	54.97	16	45	16.31	47.92
2004	53.10	16	47	17.00	47.52
2005	58.44	20	63	18.33	51.45
2006	59.95	20	59	18.69	51.35
2007	57.67	20	61	19.69	55.09
2008	53.55	21	64	21.48	63.04
2009	49.83	19	84	19.43	75.57
2010	53.37	19	79	19.35	74.73
2011	51.80	19	82	21.05	83.84
2012	50.97	19	85	21.67	93.08
2013	52.54	19	87	23.01	92.61
2014	53.70	19	82	23.17	87.58
2015	54.74	19	78	22.97	88.76
2016	57.26	19	73	23.31	86.18

Descriptive statistics for firms in *IDAN*, which contains all hirings, and *LON*, which contains information on work time. Column (2) contains the percentage of firms present in both *IDAN* and *LON*. Columns (3) and (4) contain the average number of employees in *IDAN* and *LON*. Columns (5) and (6) contain the average revenue in millions of current DKK in *IDAN* and *LON*.

A.1 Description of Work Hours

In the LON register, performed hours of work are characterized by the following equation:

$$\text{Performed work hours} = \text{Contract hours} + \text{Overtime} - \text{Paid absence} - \text{Vacation} - \text{Holiday hours}, \quad (5)$$

where "contract hours" denote the extent of a normal work week, typically being 37 hours. The variable in LON for vacation hours is missing for 2009 and after, but could be approximately imputed using the remaining hour components and the above equation. The mean for each component for the time period is given by Table A.3 below, while the distributions of performed work hours, contract hours, and paid absence for the years 2005, 2010, and 2015 are given in Figure A.4. Both Table A.3 and Figure A.4 display very little time variation in performed hours of work and its components.

Overtime defines the number of work hours beyond the normal hours in the employment contract. For an hour of overtime work, the employee can be compensated by either an hourly wage increase, or by getting an extra hour of paid absence. Work on Sundays and public holidays is also considered overtime work.

Paid absence denote the hours of work paid out to the employee due to own illness, children's illness, maternity leave, accidents or similar. Importantly, paid absence also covers the additional vacation leave including the "sixth vacation week", which is not covered by the Danish Holiday Act. This additional vacation leave can often be paid out as a wage increase instead. Since the firms' wage payouts are not affected by the absence of the employee, they may have low incentive to correctly report them. For this reason. Statistics Denmark imputes values in cases with unrealistically low absence for the employees.

Vacation is hours paid vacation time for the employee. Every employee is entitled to at least five weeks of paid leave for one calendar year under The Danish Holiday Act, as long as they have worked for a full calendar year, before the beginning of the holiday year. If that is not the case, the employee is entitled to 2.08 days of holiday leave per month of employment. For this reason, a reduction in vacation time may be seen for worker who have recently switched employers.

Holiday hours are hours of work lost due to holidays on weekdays.

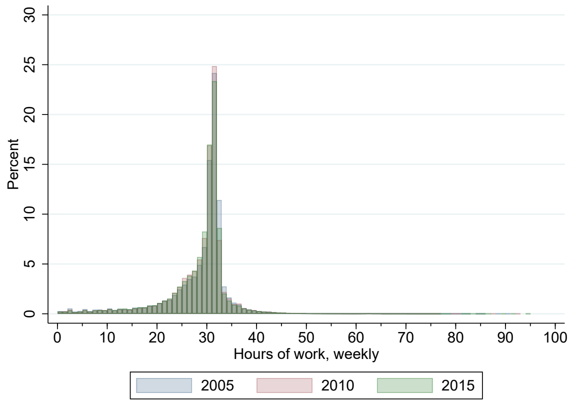
Table A.3: Mean of Types of Hours (Eq. 5) Over Time

Year	Performed work hours	Contract hours	Overtime	Paid absence	Vacation	Holiday hours
2002	29.24	34.16	0.50	1.56	2.92	0.99
2003	29.17	34.10	0.46	1.57	2.90	0.99
2004	28.65	33.70	0.42	1.68	2.90	0.89
2005	29.15	34.08	0.46	1.62	2.89	0.88
2006	29.52	34.37	0.51	1.33	2.94	1.08
2007	28.94	33.81	0.51	1.44	2.94	1.00
2008	28.75	33.68	0.46	1.46	2.92	0.99
2009	28.96	34.10	0.39	1.53	.	0.98
2010	29.05	34.48	0.41	1.68	.	1.10
2011	29.01	34.40	0.39	1.65	.	1.09
2012	28.97	34.53	0.38	1.65	.	1.25
2013	29.09	34.35	0.41	1.53	.	1.10
Total	29.04	34.15	0.44	1.56	2.92	1.03
Observations			10,719,485			

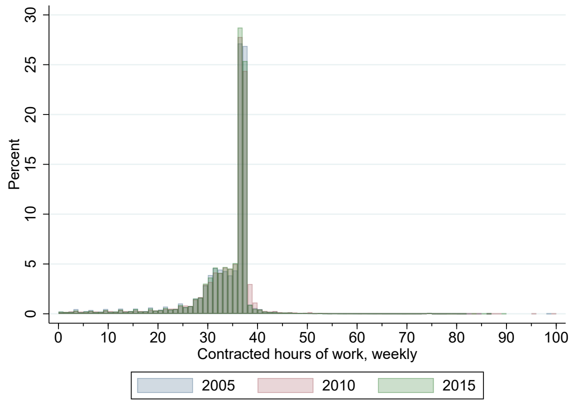
All hours are transformed from annual to approximately weekly by multiplying them by $\frac{7}{365}$. The table is made under the same conditions as the main regressions (and column (3) of Table A.1), but also imposing that contract hours are non-missing and non-zero.

Figure A.4: Distribution of Performed Work Hours, Contract Hours, and Paid Absence

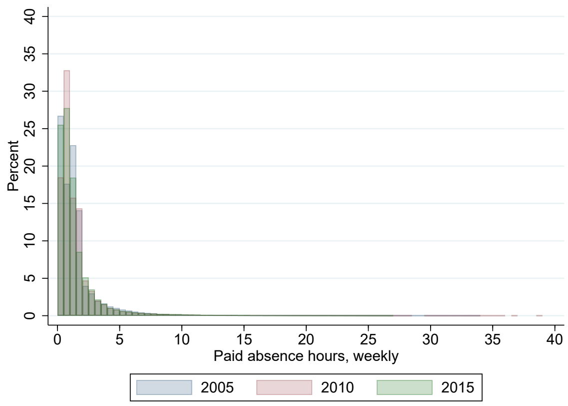
(a) Performed Work Hours



(b) Contract Hours



(c) Paid Absence

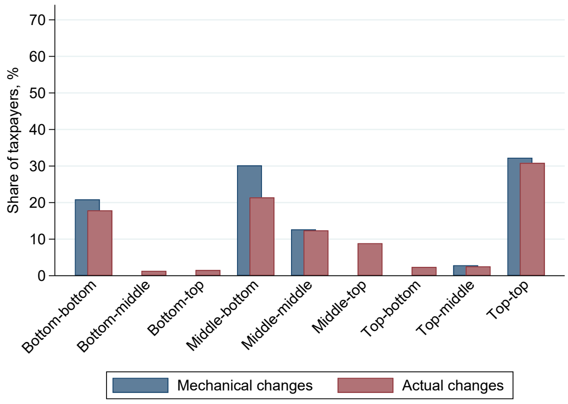


The figure shows the distribution of performed work hours, contract hours, and paid absence for the years 2005, 2010, and 2015. All hours are transformed from annual to approximately weekly by multiplying them by $\frac{7}{365}$. For Figure A.4a and Figure A.4b, the bin size is 1, while it is 0.5 for Figure A.4c. All figures are made under the same conditions as the main regressions, Table 2, and each bin is conditional on containing > 5 observations.

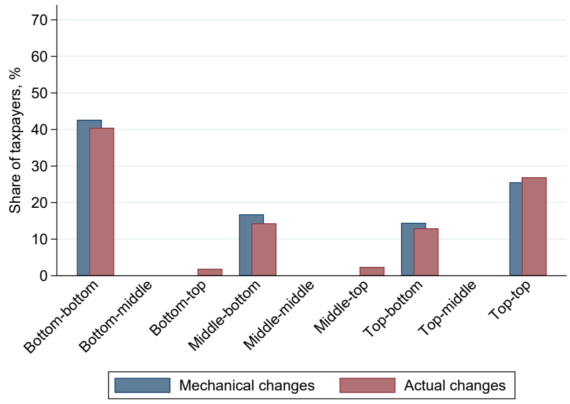
B Additional Figures and Tables

Figure B.1: Variation in Tax Bracket Compositions From Tax Reforms

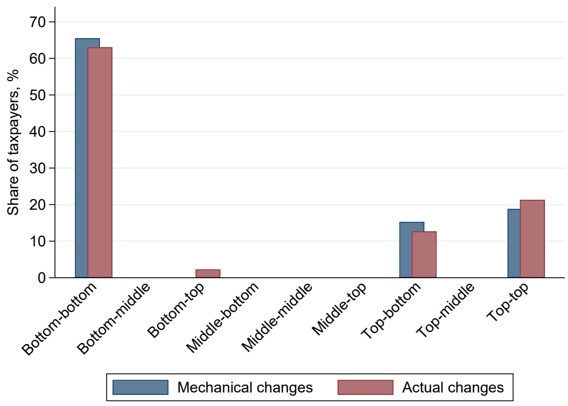
(a) 2003 Reform: 2003-06 change



(b) 2007 Reform: 2007-10 change



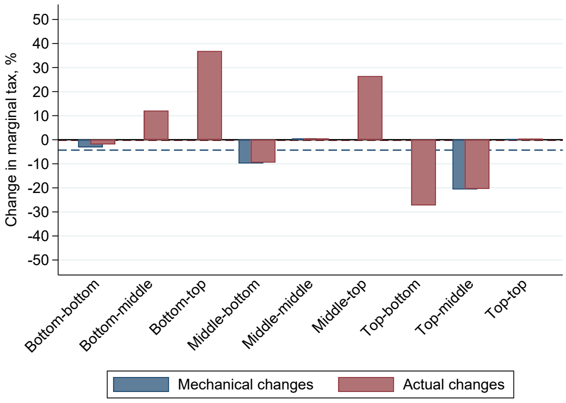
(c) 2012 Reform: 2012-15 change



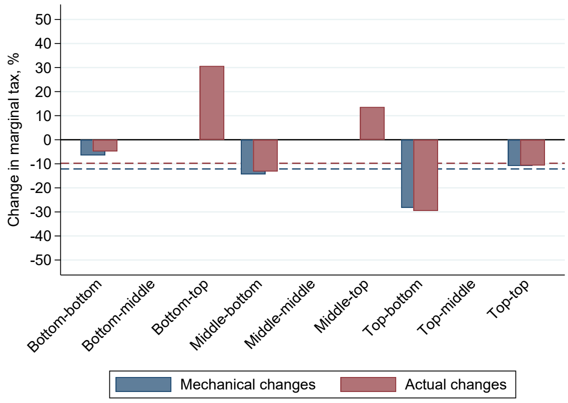
The figure shows the share of taxpayers changing from one tax bracket to another over three years. The group of taxpayer is denoted on the horizontal axis. The groups are determined based on which bracket the taxpayer belongs to before and after the three year change. E.g. in Figure B.1a, "Top-middle" refers to taxpayers who are in the top tax bracket in 2003, but in the middle bracket in 2006. "Mechanical change" refers to the share of taxpayers changing tax bracket implied by the change in tax system over time. "Actual change" refers to the actual share of taxpayers changing brackets in the three year period. Note, the 2007 reform captures changes in the tax system caused by both the 2007 and 2009 reforms.

Figure B.2: Variation in Marginal Taxes From Tax Reforms

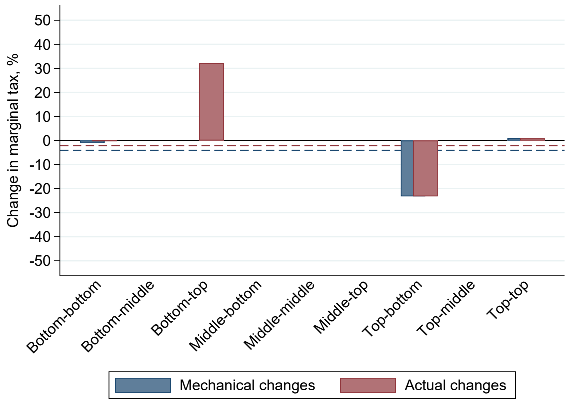
(a) 2003 Reform: 2003-06 change



(b) 2007 Reform: 2007-10 change



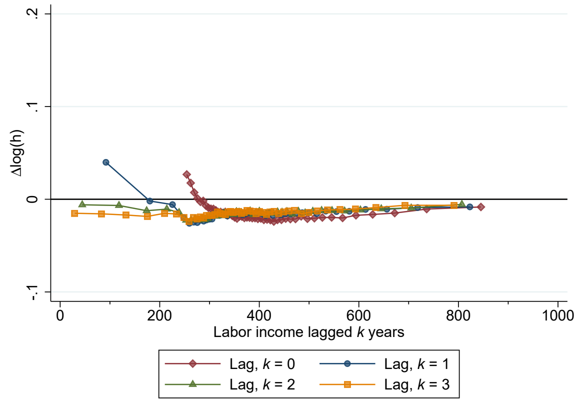
(c) 2012 Reform: 2012-15 change



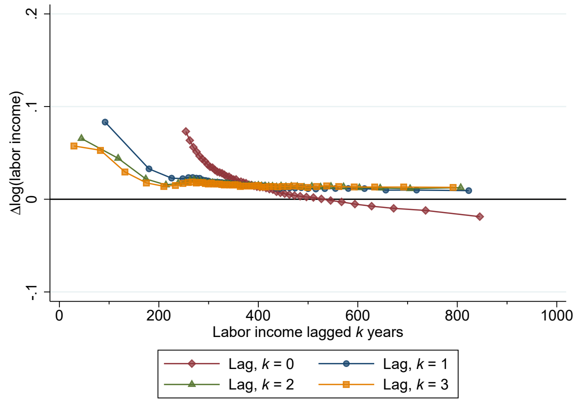
The figure shows the three-year percentage change in marginal taxes for different groups of taxpayers. The group of taxpayer is denoted on the horizontal axis. The groups are determined based on which bracket the taxpayer belongs to before and after the three year change. E.g. in Figure B.2a, "Top-middle" refers to taxpayers who are in the top tax bracket in 2003, but in the middle bracket in 2006. "Mechanical change" refers to the change in marginal tax implied by the change in tax system over time. These are calculated as described in Equation (2) with *k=0*. "Actual change" refers to the actual changes in marginal taxes for the taxpayer in the three year period. Note, the 2007 reform captures changes in the tax system caused by both the 2007 and 2009 reforms. The dashed horizontal lines indicate the mean mechanical (navy blue) and actual (maroon red) changes across all brackets.

Figure B.3: Illustration of Mean Reversion

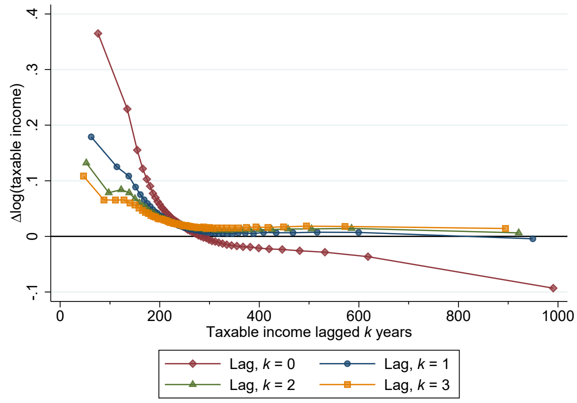
(a) Change in log Work Hours



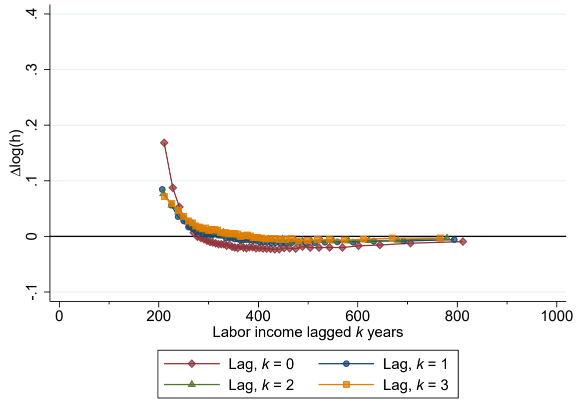
(b) Change in log Labor Income



(c) Change in log Taxable Income



(d) Change in log Work Hours (No Income Condition)



The figure shows the correlation between the three-year log change in hours of work, labor income or taxable income and with the related concept of initial income lagged *k* times. E.g., in Figure B.3a, the correlation between the three-year log change in hours of work and initial labor income lagged 0-3 years is shown. In Figure B.3d, "No Income Condition" refers to not imposing the condition of having greater or equal to 250,000 (2015 DKK) initial labor income. All four graphs are only depicted for initial labor or taxable income $\leq$ 1,000,000, while Figure B.3d also imposes labor income $\geq 250,000$ (for all lags) for visual simplicity.

Table B.1: Main Regression Results - First Stage Regressions

	(1)	(2)	(3)	(4)
	$\Delta\ln(1-\tau)$	$\Delta\ln(1-\tau)$	$\Delta\ln(1-\tau)$	$\Delta\ln(1-\tau)$
$\Delta\ln(1-\tau^p_{it,t-0})$	0.7661^{***}	0.6673^{***}		
	(0.0006)	(0.0006)		
$\Delta\ln(1-\tau^p_{it,t-2})$			0.2704^{***}	0.2372^{***}
			(0.0007)	(0.0007)
F-test statistic	1442042.1	1327014.5	145300.0	108444.8
Time fixed effects	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes
Socio-economic controls	Yes	Yes	Yes	Yes
Instrument lags	0	0	2	2
Income controls	None	10-piece spline	None	10-piece spline, lag 2
Observations	10,755,733	10,755,733	10,755,733	10,755,733

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

First stage regressions for Table 2. The independent variable is the net of marginal tax rate on additional labor income, which is regressed on the instruments described by Equation (2) with $k = 0$ for columns (1) and (2) and $k = 2$ for columns (3) and (4). Control variables are included as in Table 2.

Table B.2: Main Estimates - Different Difference Sizes

	(1) $\Delta_s \ln h$, $s = 1$	(2) $\Delta_s \ln h$, $s = 2$	(3) $\Delta_s \ln h$, $s = 3$	(4) $\Delta_s \ln h$, $s = 4$	(5) $\Delta_s \ln h$, $s = 5$	(6) $\Delta_s \ln h$, $s = 6$
$\Delta_s \ln(1-\tau)$	0.0302***	-0.0004	0.0814***	0.1267***	0.1201***	0.0760***
	(0.0101)	(0.0087)	(0.0091)	(0.0105)	(0.0115)	(0.0135)
First stage F-statistic	82618	122842	108445	82281	67213	47832
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Socio-economic controls	Yes	Yes	Yes	Yes	Yes	Yes
Instrument lags	2	2	2	2	2	2
Income controls	10-piece spline, lag 2	10-piece spline, lag 2	10-piece spline, lag 2	10-piece spline, lag 2	10-piece spline, lag 2	10-piece spline, lag 2
Observations	11,997,475	11,252,189	10,755,733	9,290,833	7,999,653	6,807,745

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

Labor supply elasticities estimated as in column (4) of Table 2, but with differing difference sizes, s, given in the column header. The dependent variable is the s-year log change in *performed* hours of work and the independent variable of interest is the s-year log change in the marginal tax on additional labor income. This has been instrumented by the mechanical changes, calculated by Equation (2), where k is given by the row "Instrument lags". The row "Income controls" denotes which income controls are included. "10-piece spline" refers to a 10-piece spline with breakpoints at deciles of log labor income. All specifications include time, individual fixed effects, socio-economic controls, and municipality fixed effects. The socio-economic controls are as listed in Section 3.

Table B.3: Main Estimates - Different Lag Lengths

	(1)	(2)	(3)	(4)	(5)
	$\Delta \ln h$	$\Delta \ln h$	$\Delta \ln h$	$\Delta \ln h$	$\Delta \ln h$
$\Delta \ln(1-\tau)$	0.0241***	0.0418***	0.0814***	0.1162***	0.1536***
	(0.0032)	(0.0071)	(0.0091)	(0.0118)	(0.0144)
First stage F-statistic	1327015	173499	108445	65587	44912
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes
Socio-economic controls	Yes	Yes	Yes	Yes	Yes
Instrument lags	0	1	2	3	4
Income controls	10-piece spline, lag 0	10-piece spline, lag 1	10-piece spline, lag 2	10-piece spline, lag 3	10-piece spline, lag 4
Observations	10,755,733	10,728,780	10,755,733	10,609,874	10,507,236

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

Labor supply elasticities estimated as in column (4) of Table 2, but with differing lag lengths, k. The dependent variable is the three-year log change in *performed* hours of work and the independent variable of interest is the three-year log change in the marginal tax on additional labor income. This has been instrumented by the mechanical changes, calculated by Equation (2), where k is given by the row "Instrument lags". The row "Income controls" denotes which income controls are included. "10-piece spline" refers to a 10-piece spline with breakpoints at deciles of log labor income. All specifications include time, individual fixed effects, socio-economic controls, and municipality fixed effects. The socio-economic controls are as listed in Section 3.

Table B.4: Main Estimates - Importance of Income Controls

	(1) log labor income	(2) 5-piece spline	(3) 10-piece spline	(4) 20-piece spline	(5) 10-piece spline and log deviation	(6) 10 piece spline and 10-piece spline of log deviation
$\Delta\ln(1-\tau)$	0.0421***	0.0756***	0.0814***	0.0854***	0.0865***	0.0866***
	(0.0080)	(0.0091)	(0.0091)	(0.0091)	(0.0091)	(0.0090)
First stage F-statistic	138248	108819	108445	108545	109210	109211
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Socio-economic controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,755,733	10,755,733	10,755,733	10,755,733	10,755,733	10,755,733

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

Labor supply elasticities estimated by 2SLS regressions including various income control specifications. The dependent variable in all specifications is the three-year log change in performed hours of work. The independent variable of interest is in all cases the three-year log change in the marginal tax on additional labor income. This has been instrumented by the mechanical changes, calculated by Equation (2), where *k=2*. The column header denotes which specification of income controls are used. In all cases, these are based on log labor income lagged 2 years. "log-deviation" refers to the log of the absolute value of the deviation between log labor income in year *t-2* and *t*. All specifications include time and individual fixed effects, socio-economic controls, and municipality fixed effects. The socio-economic controls are as listed in Section 3.

Table B.5: Main Estimates - Sensitivity to Inclusion of Control Variables

	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\ln(1-\tau)$	0.0613***	0.0791***	0.0822***	0.0814***	0.0730***	0.0857***
	(0.0091)	(0.0091)	(0.0091)	(0.0091)	(0.0181)	(0.0090)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Human capital controls	No	Yes	Yes	Yes	Yes	Yes
Social & Industry controls	No	No	Yes	Yes	Yes	Yes
Municipality controls	No	No	No	Yes	Yes	Yes
Year × income controls	No	No	No	No	Yes	No
"Residualized" instrument	No	No	No	No	No	Yes
Observations	10,755,733	10,755,733	10,755,733	10,755,733	10,755,733	10,755,733

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

Labor supply elasticities estimated as in column (4) of Table 2, but with varying levels of control variables. "Human capital controls" denotes controls for age, experience, and education. "Social & Industry controls" denotes dummies for being married, age of youngest child, and workplace industry. "Year × income controls" indicates interacting year dummies with each of the 10-piece spline segments of initial log labor income lagged 2 years. "Residualuzed instrument" indicates using the residuals from regressing the instrument variable on a 10-piece spline of log labor income lagged 2 years, which replaces the standard instrument and income controls. Standard errors have not been corrected for this "first stage"-type regression in column (6).

Table B.6: Empirically Testing the Validity of Instruments

	(1)	(2)	(3)	(4)
	$\Delta\ln h$	$\Delta\ln h$	$\Delta\ln h$	$\Delta\ln h$
$\Delta\ln(1-\tau)$	-0.0179***	0.0238***	0.0324***	0.0904***
	(0.0027)	(0.0032)	(0.0070)	(0.0080)
First stage F-statistic	505066	462397	62590	49280
Diff-in-Sargan P-value	0.180	0.000	0.000	0.300
Time fixed effects	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes
Socio-economic controls	Yes	Yes	Yes	Yes
Instrument lags	0, 1, 2	0, 1, 2	2, 3, 4	2, 3, 4
Income controls	None	10-piece spline, lag 0	None	10-piece spline, lag 2
Observations	10,666,320	10,666,320	10,666,320	10,666,320

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered by individual in parentheses.

Labor supply elasticities estimated by 2SLS regressions. The dependent variable in all specifications is the three-year log change in hours of work. The independent variable of interest is in all cases the three-year log change in the marginal tax on additional labor income. This has been instrumented by the mechanical changes, calculated by Equation (2), where k is given by the row "Instrument lags". The row "Income controls" denotes which income controls are included. The Difference-in-Sargan P-value has been calculated as described in Section 6.4.

Economics Working Papers

2021-03 Chang Liu, Tor Eriksson and Fujin Yi: Offspring Migration and Nutritional Status of Left-behind Older Adults in Rural China

2021-04 Benjamin Lochner and Bastian Schulz: Firm Productivity, Wages, and Sorting

2021-05 Giovanni Pellegrino, Efrem Castelnuovo and Giovanni Caggiano: Uncertainty and Monetary Policy during the Great Recession

2021-06 Anna Folke Larsen and Marianne Simonsen: Social emotional learning in the classroom: One-year follow-up results from PERSPEKT 2.0

2021-07 Johannes Schünemann, Holger Strulik and Timo Trimborn: Optimal Demand for Medical and Long-Term Care

2021-08 Tom Lane, Daniele Nosenzo and Silvia Sonderegger: Law and Norms: Empirical Evidence

2021-09 Ina C. Jäkel: Export Credit Guarantees: Direct Effects on the Treated and Spillovers to their Suppliers

2021-10 Martin Paldam: Measuring Democracy - Eight indices: Polity, Freedom House and V-Dem

2021-11 Nicola Maaser and Thomas Stratmann: Costly Voting in Weighted Committees: The case of moral costs

2021-12 Martin M. Andreasen, Giovanni Caggiano, Efrem Castelnuovo and Giovanni Pellegrino: Why Does Risk Matter More in Recessions than in Expansions?

2021-13 Søren Albeck Nielsen: How to Cope with Dyslexia: The Effects of Special Education on Academic Performance, Personality Traits, and Well-being

2022-01 Michael Koch and Ilya Manuylov: Measuring the Technological Bias of Robot Adoption and its Implications for the Aggregate Labor Share

2022-02 Olga Balakina, Claes Bäckman, Andreas Hackethal, Tobin Hanspal and Dominique M. Lammer: Good Peers, Good Apples? Peer Effects in Portfolio Quality

2022-03 Kazuhiko Sumiya and Jesper Bagger: Income Taxes, Gross Hourly Wages, and the Anatomy of Behavioral Responses: Evidence from a Danish Tax Reform

2022-04 Hans Schytte Sigaard: Labor Supply Responsiveness to Tax Reforms