

THE QUALITY OF THE ESTIMATORS OF THE ETI

THOMAS ARONSSON, KATHARINA JENDERNY, AND GAUTHIER LANOT

ABSTRACT. Measuring the elasticity of taxable income (ETI) is central for tax policy design. Yet, there are few arguments which support or infirm that current methods yield measurements of the ETI that can be trusted. Our first purpose is to use simulation methods to assess the bias and precision of the prevalent methods used in the literature (IV estimation and bunching methods). Thereby, we aim at (i) explaining the huge differences in empirical results, and (ii) providing arguments in favor of or against using these methods. Our second purpose is to suggest indirect inference estimation to improve the quality of the measurement.

We find that the IV regression estimators may suffer from considerable bias and be quite imprecise, whereas the bunching estimators perform better in our controlled environment. We also show that using more of the information available in the data, estimators based on indirect inference principles produce more precise estimates of the ETI than any of the most commonly used methods.

JEL Classification: H24; H31; D60

Keywords: Elasticity of Taxable Income, Income Tax, Indirect Inference, IV estimation, Bunching, Monte Carlo simulations

1. INTRODUCTION AND LITERATURE

Background and general purpose. The elasticity of taxable income with respect to the marginal net-of-tax rate (ETI) is a central statistic for tax policy design.¹ In fact, it is often referred to as “sufficient statistic”, i.e. a parameter that provides sufficient information on the behavioral response to marginal taxation (Feldstein, 1995). Yet, although measurable in principle, there is little agreement on the empirical size of the ETI. The two dominating methodological approaches, the instrumental variable (IV) regression-based approach and the

DEPARTMENT OF ECONOMICS, UMEÅ UNIVERSITY, SWEDEN

E-mail addresses: thomas.aronsson@umu.se, katharina.jenderny@umu.se, gauthier.lanot@umu.se.

Acknowledgements: We would like to thank Frank Fossen, David Granlund, and Magnus Wikström as well as the participants of the ZEW public finance conference 2017 and the IIPF conference 2017 for helpful comments and suggestions. Financial support from Handelsbanken Foundation (project P2016-0140:1) is gratefully acknowledged.

¹The marginal net-of-tax rate is defined as one minus the marginal tax rate.

bunching approach, fail to produce similar point estimates for the ETI. Point estimates based on the regression approach typically exceed point estimates based on the bunching approach, sometimes by an order of magnitude, even using the same data. Furthermore, the performance of these methods in terms of bias and precision is unclear. Despite a large body of literature, policy makers are thus still largely in the dark when designing the tax system.

Our purpose in this paper is to use simulation methods to assess the bias and precision of the prevalent methods used in the literature. Thereby, we aim at explaining the differences in empirical results, and providing arguments in favor of or against using these methods. Furthermore, we suggest to measure the ETI with an indirect inference estimator.

Our expectations are based on the reasoning that, given the specification of the earnings function (and in the absence of information on the hourly wage), all observations and all earnings histories are informative about the ETI. Yet, both the bunching approach and regression-based methods use only a fraction of the available information. The regression approach typically focuses on differences and disregards the information on bunching, while the bunching approach only uses the cross-sectional information of observations near the kink points. The indirect inference method, by contrast, uses both the cross sectional information and earnings histories. We therefore expect the suggested method to perform better.

In order to assess the bias and precision of the bunching method, regression-based methods, and the indirect inference method, we simulate data according to a well-specified economic model and empirically based assumptions concerning its parameters (such as the autocorrelation process in incomes). We then set different "true" values for the ETI and analyze the ability of the different estimators to reproduce the true parameter estimate in a Monte Carlo study. We also assess the performance of the estimators depending on the tax environment, such as the variability of tax rates over time.

Our simulations show that the indirect inference estimator is considerably more precise than both the bunching and IV estimators under a variety of conditions. Our analysis helps to reconcile existing results and provides guidance on which methods to choose, eventually resulting in more credible estimates of the ETI.

Literature and Issues. Let us now turn to the two prevailing estimation methods of the ETI, the regression-based method and the bunching method. We discuss the methods in more detail in the context of our model specification in section 4.

The regression approach seeks to identify the ETI by comparing relative changes in taxable income of tax units or groups of tax units between two periods to relative changes in their net-of-tax rates. In the early literature, group-based comparisons of income shares were conducted on cross-sectional data (Feenberg and Poterba, 1993, see also the review by Saez et al., 2012). Feldstein, (1995) was the first study to use panel data, which allowed to keep the composition of groups of tax units constant over time to avoid endogenous selection into groups. Later studies

used difference-in-difference regressions including further control variables instead of simple group mean comparisons.

The regression-based approach faces two main methodological challenges. First, following tax units over time introduces the problem of mean reversion, which is akin to an initial conditions problem, and describes the correlation between the error term and the dependent variable in first differences. A high-income tax unit in the initial period is likely to have a lower income in the following period for idiosyncratic reasons, irrespective of tax rate changes. The mean reversion problem is typically addressed by controlling for the initial income level.² Second, the marginal net-of-tax rate is an endogenous variable to the change of the income level in directly progressive tax systems, and instrumental variables techniques are used to account for this endogeneity. A common instrument for the change in the net-of-tax rate is a hypothetical change in the net-of-tax rate that uses the tax schedules of the period for which the ETI is to be measured, but applies them only to the start-year income (Auten and Carroll, 1999; Carroll, 1998; Gruber and Saez, 2002). Weber, (2014) argued that using the initial income level for the instrumentation does not solve the endogeneity problem, and suggests to use higher-order lags of the income instead.

Regression-based estimates of the ETI differ greatly across methods and countries (see, for example, Neisser, 2017, who shows that estimates range roughly between -1.5 and 2). Some of these differences can be explained by conceptual differences, such as different tax systems and taxpayer types. For instance, the ETI is expected to be higher when the share of self-reported income is large, and if deduction possibilities are abundant (Doerrenberg et al., 2017; Kleven and Schultz, 2014; Kopczuk, 2005). Other differences reflect methodological advances, such as the proper instrumentation of endogenous variables. The approach of Feldstein, (1995) does not use any instrumentation to account for the endogeneity of the net-of-tax rate, and yields large elasticity estimates, ranging roughly between 1 and 3. Using US data and controlling for lagged income, Gruber and Saez, (2002) obtain elasticities between 0.4 and 0.6, depending on the income control. Weber, (2014) shows that instrumentation with longer time lags of the instruments leads to baseline estimates of the ETI that are twice as large as those found by Gruber and Saez, (2002), using the same data.

The bunching approach provides an alternative method to estimate the ETI. Modern income tax and benefit systems are usually piecewise linear: the marginal tax rate is typically constant within well defined intervals of the taxable income, while it changes in a discontinuous manner between intervals. This creates either kinks or notches to the budget constraint of the tax payer. At a kink point, the marginal tax rate changes in a discontinuous fashion (e.g., between two proportional tax brackets), while the tax due changes in a discontinuous fashion at a

²As we will argue in part 4, it is not clear that this control variable will solve the missing variable problem.

notch. In practice, notches occur less frequently than kink points. We thus focus on kink points when discussing the methodological approach.³

Kink points help identifying the ETI based on the behavior of agents located at or close to the kink points. Household tax positions will typically bunch at any kink of the tax schedule. Optimizing individuals choose taxable income so as to equalize their marginal rate of substitution (MRS) between the utility cost of acquiring an additional unit of income (e.g., in terms of leisure foregone) and the disposable income to the marginal net-of-tax rate. At a kink point, economic theory predicts that all tax units whose MRS is less than or equal to the net-of-tax rate in the first, but not the second bracket should generate taxable income up to that kink point exactly. The amount of tax units at the kink point can therefore identify the average responsiveness of tax units to the marginal tax rate (i.e., the ETI), by contrasting the excess mass of bunching tax units to a counter-factual income distribution in the absence of a kink point. Bastani and Selin, (2014), Chetty et al., (2011), Hargaden, (2015), Kleven and Waseem, (2013), and Saez, (2010) use the amount of bunching relative to a theoretical counter-factual without bunching, to measure the behavioral response to the tax code. Kleven, (2016) provides a recent methodological review.

Similar to the regression approach, the bunching approach to measuring the ETI faces methodological challenges. The first challenge is that the construction of a counter-factual income distribution in the absence of a kink required to operationalize the measurement is not straightforward, and its functional form matters for the result. Furthermore, as the distribution of taxable income depends on the tax rate, the actual distribution above the kink is not suitable to fit a counter-factual, as it is affected by the tax rate change even if tax units choose higher incomes than the kink income.⁴ The second challenge is that the bunching typically occurs not only at the exact kink income, but in an interval, which has to be specified by the researcher. The third challenge is that the estimator is local by definition, and its validity is restricted to a particular income level, and possibly household type at that level. The fourth challenge is that optimization frictions are likely to bias the estimates downwards, especially in the case of wage earners (Bastani and Selin, 2014; Chetty et al., 2013).

Both regression and bunching estimators only use a limited amount of the data available. Regression methods are typically based on a linearization of the budget constraint and do not take into account that tax kinks have an influence on behavior, while bunching methods use cross-sectional information only.⁵ By

³For discussions on notches, see Hargaden, (2015), Kleven and Waseem, (2013), Kleven, (2016), and Slemrod, (2013), for example.

⁴Chetty et al., (2013) evade using a functional form, and instead use regional variation in information on the EITC schedule to construct counter-factual income distributions.

⁵An exception in the context of health spendings are Einav et al., (2015, 2017), who allow for inter-temporal substitution of spendings.

comparison, many earlier studies on the labor supply (see the reviews by Blundell and Macurdy, 1999 and Blundell et al., 2007) used maximum likelihood (ML) methods to estimate jointly the behavioral parameters and the parameters that describe the distribution of heterogeneity in the population. In practice, while ML methods are suitable to overcome the limitations of the regression and bunching estimators in general, they are analytically difficult to apply in contexts where the analyst wishes to account for repeated observations over time or/and when the distribution of the unobserved components of the model is not normal. Modern simulation based methods provide a possible alternative by combining the advantages of ML estimation with mathematical feasibility. We suggest to use indirect inference principles (see Gouriéroux et al., 1993 for details) to estimate the ETI.

The indirect inference approach has been applied in different contexts (see Browning et al., 2010, Low and Pistaferri, 2015, and Nenov, 2015, for example), but not in the literature measuring the ETI. It relies on two elements; first, a model which potentially generates the data but depends on a set of unknown parameters, among those the ETI, and, second, a large enough set of auxiliary statistics which can be estimated on the sample data as well as on the simulated data from the model for any value of the parameters. Assuming the theoretical model is a good description of the process that generates the observed data, the values the auxiliary statistics take when measured from the observed data will be similar to the values of the same statistics when measured from the simulated data at the correct parameters. Gouriéroux et al., (1993) show that the parameter values that minimize the distance between the estimated auxiliary statistics obtained from the sample data and the ones obtained from the simulated data will have good statistical properties and in particular (asymptotic) unbiasedness.

The empirical evidence is captured by the measurement on the observed sample data of auxiliary statistics. The inference is indirect since the choice of parameter estimates for the model is guided by the ability of the simulated data to generate auxiliary statistics values that are close (or identical) to the ones the empirical data generates. The method can be applied in any context where it is (relatively) easier to simulate data from a given theoretical model than it is to calculate the moments or the likelihood the theoretical model implies. We argue that this is in general the case in the context of the estimation of the ETI.

Contribution. Even accounting for conceptual, methodological and data differences, the literature has not been able to narrow down the plausible size of the parameter even by order of magnitude. As we indicate above, point estimates of the ETI are also different between the regression and the bunching approach, with bunching results being arguably smaller.⁶ This raises concerns about the reliability of the estimates obtained.

⁶For example, Kleven and Schultz, (2014) report that their estimates obtained from a large Danish tax reform is an order of magnitude larger than the result obtained by Chetty et al., (2011) using a single large Danish kink. We report some illustrative estimates from the literature in Tables 18 and 19 in Appendix B.

In the following, we will use the term “conventional methods” when referring to the regression and bunching methods of estimating the ETI. One contribution of our study is to assess the performance of these conventional methods, in terms of bias and precision, using Monte Carlo simulation techniques. Another is to use indirect inference estimators, which are novel in the context of ETI estimation, and compare their performance with the performance of the conventional estimators.

We find that the IV regression estimators may suffer from considerable bias and be quite imprecise, whereas the bunching estimators perform better in our controlled environment. We also show that using more of the information available in the data, estimators based on indirect inference principles produce more precise estimates of the ETI than any of the conventional methods.

The paper is structured as follows: In part 2, we set up a simple behavioral model, which we use to simulate data as described in part 3. In part 4, we describe the estimators, both conventional and newly suggested, that we apply to our simulated data. In part 5 we present and discuss the results. Concluding remarks are presented in part 6.

2. BEHAVIORAL MODEL

General framework. In order to simulate data generated through optimization behavior of individuals, we need a behavioral model that describes the choice of taxable income in response to the net-of-tax rate. The general framework we present here is designed to provide a simple model in which the main methodological questions are nevertheless relevant. For this reason, we focus on labor income and assume that individuals respond to the wage they are offered. In response to an offered wage, individuals determine their level of work effort, and together the offered wage and the labor supply determine a given individual’s gross earnings.

The model presented below assumes that labor is the only income source, and that the utility is quasi-linear in private consumption. The reason for these simplifications is that we aim to define the most stylized model framework that will still reflect the core estimation issues in the literature. Both assumptions can of course be relaxed in principle. In fact, we argue in part 4 that the estimator we propose is more suited to relax these assumptions than the methods currently used in the ETI literature. Our exact specification is the one adopted by Saez, (2010), and therefore corresponds exactly to the seminal bunching approach. As we show below, the model also reproduces the general specification used in the regression-based approach, where it is more common to state the reduced-form relationship between the change in income and the change in the net-of-tax rate directly than to formulate a complete structural model (see, e.g., Gruber and Saez, 2002). Consequently, our model provides a favorable framework for the estimators we analyze.

The specification rests on the preferences which yield the basic log linear specification of the labor supply function

$$u(c, h) = c - \gamma \frac{\eta}{1 + 1/\alpha} \left(\frac{h}{\eta}\right)^{1 + \frac{1}{\alpha}}, \quad (1)$$

where α , γ and η are all positive.⁷ The labor supply function takes the form:

$$\ln h^*(w, \eta) = -\alpha \ln \gamma + \alpha \ln w + \ln \eta. \quad (2)$$

Equation (2) gives a direct interpretation to the parameters of the utility function: α is the wage elasticity of the labor supply, and both η and γ describe the disutility of work. γ determines the average disutility of work, while η is assumed to be one on average and introduces heterogeneity between individuals. $\eta > 1$ corresponds to a below-average disutility of work, while $\eta < 1$ corresponds to an above-average disutility of work. Observe that the specification excludes income effects. Chetty, (2012) suggests that α (the labor supply elasticity) = 0.33 is a credible central prior based on his reading of the accumulated (US) evidence on the intensive margin.

The log linear specification has another appealing feature in describing a straightforward relationship between the offered wage, w , and gross earnings/taxable income wh . Indeed we have,

$$\ln wh^* = -\alpha \ln \gamma + (\alpha + 1) \ln w + \ln \eta, \quad (3)$$

where $\alpha + 1$ is now the elasticity of earnings relative to the wage.

In the following, we adjust the model in two dimensions that are relevant to most empirical applications. First, we account for the fact that we typically do not observe the wage and the disutility of work, but only earned income. We show that we can treat the variability of both unknown factors as one single unknown component in this case. Second, we introduce taxation in the model, which enables us to define the elasticity of taxable income and to address the endogeneity of the net-of-tax rate to the optimal choice of earnings. The latter is crucial for the empirical estimation using the regression approach.

The variability of the wage and the disutility of work. Heterogeneity between agents in this model stems from the variability of the wage offer and the variability of the disutility of work. Yet, we only observe earned income which, in turn, depends on both w and η . It is therefore useful to understand the possible interactions between the two. Observe that it is always possible to rewrite the preferences over a bundle (c, h) given the disutility of work η (as defined in Equation (1)), as the utility over a bundle (c, wh) given the wage w and the disutility of work η , i.e.,

$$u(c, h) = c - \gamma \frac{\eta}{1 + 1/\alpha} \left(\frac{wh}{w\eta}\right)^{1 + \frac{1}{\alpha}}, \quad (4)$$

⁷Saez, (2010) does not specify γ as he does not assume η to be one on average.

and we deduce the preference over consumption and earnings,

$$v(c, wh) = c - \gamma \frac{1}{1 + 1/\alpha} (wh)^{1 + \frac{1}{\alpha}} (w^{\alpha+1} \eta)^{-\frac{1}{\alpha}}. \quad (5)$$

When specified in this fashion the disutility of work depends on the quantity

$$\omega \equiv w^{\alpha+1} \eta.$$

The behavioral assumption requires then that the worker determines earnings, wh , so as to maximize $v(c, wh)$ such that $c = wh + R$, with R being unearned income. Optimal earnings are then:

$$\ln wh^* = -\alpha \ln \gamma + \ln \omega. \quad (6)$$

This property suggests that the marginal distribution of earnings depends on the distribution of ω only. In the absence of any information about the individual wage, the distribution of optimal earnings will only be informative about the distribution of ω overall and not about the distribution of its individual components w and η . In our simulations, we can thus reduce heterogeneity to the combined unknown component ω .

Taxation. In order to use the model to predict an individual's change in optimal earnings in response to a tax rate change, we need to explicitly model income taxation. Assume that the individual's optimal labor supply choice is on a regular part of the the budget constraint (i.e., not situated at a kink or a discontinuity). It must then satisfy:

$$\ln h^* = -\alpha \ln \gamma + \alpha \ln (\tau^c[wh^*, \mathbf{x}]w) + \ln \eta, \quad (7)$$

where the amount of tax paid, $T(wh, \mathbf{x})$, depends on the level of earnings as well as on other variables observed or unobserved, \mathbf{x} . The variable $\tau^c[wh, \mathbf{x}] \equiv 1 - \tau[wh, \mathbf{x}] = 1 - \frac{\partial T}{\partial wh}$ is the marginal net-of-tax rate, defined as one minus the marginal tax rate.

We can re-formulate Equation (7) to describe the optimal level of earnings, which takes the form:

$$\ln wh^* = \kappa + \alpha \ln \tau^c[wh^*, \mathbf{x}] + \ln \omega \quad (8)$$

with $\kappa \equiv -\alpha \ln \gamma$. α measures the response of earnings to a marginal increase in the net-of-tax rate, i.e., the ETI. The expression in Equation (8) sets a starting point for the methodological discussion concerning the estimation of α where $\ln \omega \equiv (\alpha + 1) \ln w + \ln \eta$ plays the role of the unobserved component.

The specification developed in Equation (8) is the basis for most empirical measurements of the ETI. It is the simplest framework for understanding the effect of a tax system on the distribution of earnings. In most cases it is possible to trace back the empirical specification to our theoretical specification, and in particular

to interpret the parameter of the net-of-tax rate as the ETI. When using bunching methods, information from a single cross section is in principle sufficient to obtain a measurement of the ETI within this exact modeling framework.

Equation (8) is only apparently linear in the unobserved component, since the unobserved component $\ln \omega$ determines the level of earnings and, in turn, the marginal net-of-tax rate, $\tau^c[wh, \mathbf{x}]$. In the presence of a complex tax system where the marginal tax rate varies with earnings, the relationship between the wage and earnings is no longer linear. Figure 1 illustrates this fact in a simple piecewise linear case and in a smooth case. The net-of-tax rate, apparently a regressor in Equation (8), is endogenous, i.e., optimal earnings and the net-of-tax rate are determined jointly.

This endogeneity is crucial to the IV regression approach to measuring the ETI. In the economic model outlined in (8) the unobserved component and the identity of the tax bracket are correlated by construction, which requires the net-of-tax rate to be instrumented.

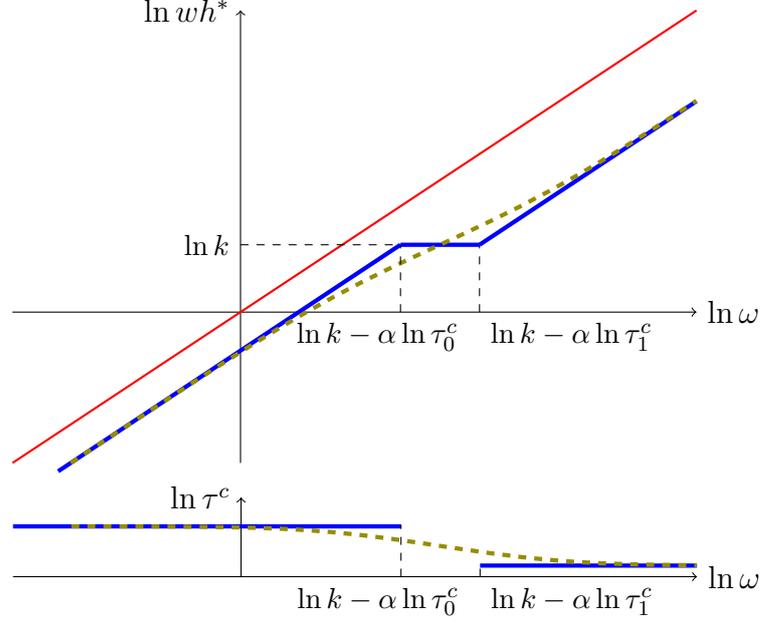
With the availability of longitudinal tax registers, the measurement of the ETI can take advantage of panel data techniques. The structure of the unobserved component, $\ln \omega$, is then further specified to capture the variability of the data within and between individual tax units over time. $\ln \omega$ is often represented as the sum of a permanent component, which reflects permanent differences between tax units, and a transitory component, which reflects differences over time within a given tax unit. Longitudinal data provides additional sources of potential instruments to account for the endogeneity of the net-of-tax rate, and allows the researchers to "remove" (difference away) the permanent differences between individual tax units.

3. SIMULATION STUDY

We use the model set out above to simulate data, which are then used to analyze the bias and precision of the estimators used in the literature. We therefore need to be more specific about the values the parameters take, in particular the ETI, and the structure of the unobserved components. We further consider the characteristics of the tax environment and their implications for the performance of the estimators.

Based on the model set out above, we focus on a design for our simulation experiments that is reduced to the essential features. In each experiment we simulate the earnings history for a sample of individuals facing a tax system with a single kink. Individuals are assumed to behave in every period according to the static model we presented above, i.e., we simulate our synthetic data according to the model that is described in Equation (8). We generate a panel of the optimal choice of 10000 individuals per year for 12 years. For each realization of the simulated data we apply a variety of estimators (bunching, IV, and Indirect Inference

FIGURE 1. Optimal earnings and the unobserved component, smooth and piecewise linear tax systems



Note: The figure illustrates the effect of the tax system on the relationship between the unobserved component, optimal earnings, and the net-of-tax rate. In a single-kinked progressive tax system (with a kink at k and marginal net-of-tax rates τ_0^c and τ_1^c below and above the kink respectively), the unobserved component $\ln \omega$ is positively correlated with optimal earnings $\ln wh^*$, and negatively correlated with the marginal net-of-tax rate.

In the upper graph, the upper (thinner) line indicates the relationship between optimal earnings and the unobserved component in the absence of taxation. The piecewise linear relationship (bold) corresponds to the same relationship with a piecewise linear tax system, while the dashed line describes the relationship the transition between the marginal net-of-tax rates is continuous.

The lower graph shows how the marginal net-of-tax rate responds to changes of the unobserved component. Over the interval $(\ln k - \alpha \ln \tau_0^c, \ln k - \alpha \ln \tau_1^c)$ optimal earnings are exactly k .

estimators) to obtain estimates of the ETI.⁸ We repeat this process a thousand times.

In order to ensure the robustness of our results, we vary the framework of our simulation in several dimensions: We consider three different "true" values for α , the ETI: 0.3, 0.6, and 0.9. Furthermore, we vary both the structure of the

⁸We describe these estimators in detail in part 4.

unobserved component ω and the tax environment. In the following we describe the latter two dimensions in more detail.

Structure of the unobserved component. We decompose the unobserved component, $\ln \omega_{it}$, into permanent and transitory terms. Furthermore, the structure of the unobserved component is then modulated in terms of the share ϕ of its variance that is attributed to the permanent effect. In summary we have

$$\ln \omega_{it} \equiv \sigma_\omega \phi u_i^P + \sigma_\omega \sqrt{1 - \phi^2} u_{it}^T \equiv \sigma^P u_i^P + \sigma^T u_{it}^T, \quad (9)$$

where u_i^P is the permanent component for individual i , and u_{it}^T is the transitory shock which affects individual i at time t . To simplify the notations we set $\sigma^P \equiv \sigma_\omega \phi$ and $\sigma^T \equiv \sigma_\omega \sqrt{1 - \phi^2}$.

In the simulations we assume that independent realizations of u_i^P are drawn from a standard normal distribution. The transitory component can take two alternative forms: First, following Weber, (2014) we assume that the transitory component follows an AR(1) process with autocorrelation ρ

$$u_{it}^T = \rho u_{it-1}^T + \lambda_{it},$$

where λ_{it} is an innovation. We set the variance of the transitory shock u_{it}^T equal to 1, i.e so that: $\mathbb{V}[\lambda_{it}] = 1 - \rho^2$. Alternatively, the transitory shock can take the form of a MA(1) process with parameter θ , so that

$$u_{it}^T = \xi_{it} + \theta \xi_{it-1},$$

where ξ_{it} is an innovation. Again we set the variance of the innovation so that the variance of the transitory component is equal to 1: $\mathbb{V}[\xi_{it}] = \frac{1}{1 + \theta^2}$.

In this way, by varying the value of ϕ and depending on the precise form of dependence of the transitory term, we can vary the stochastic structure of the unobserved component, while keeping its overall mean and variance constant. The two models imply that the autocorrelation of earnings growth is negative when the transitory processes are stationary. In the case of the AR(1) model the first autocorrelation is $\text{corr}[\Delta \ln y_t, \Delta \ln y_{t-1}] = \frac{\rho - 1}{2}$, while for the MA(1) specification the first autocorrelation takes the form $\text{corr}[\Delta \ln y_t, \Delta \ln y_{t-1}] = -\frac{(\theta - 1)^2}{1 + \theta^2 + (\theta - 1)^2}$. These expressions suggest that we can deduce realistic values for either parameters from the observed characteristics of the dynamics of earnings. Estimated on Swedish registers, the raw earnings growth first autocorrelation is about -0.2 .⁹ Hence, we set $\rho = 0.6$ when we consider an AR model and $\theta = 0.45$ when we consider instead an MA model.

⁹We obtained these figures using panel register data on Swedish labor income (ASTRID) from 2002 to 2013. The second autocorrelation is about -0.06 and the autocorrelations with longer lags take values one order of magnitude smaller in absolute value. Therefore, we focus on the first autocorrelation. Browning and Ejrnæs, (2013) report a broadly comparable autocorrelation pattern using Danish register data.

While the two models can be made to yield similar behavior for optimal earnings growth, they are distinct. The AR(1) model for the transitory shock in earnings levels implies that earnings growth satisfies another AR(1) model with the same autocorrelation (and the random component of this AR model can no longer be understood as an innovation since by construction it is correlated with the lagged dependent variable). The model of earnings growth when the transitory component is an MA(1) process becomes an MA(2) (with specific constraints on the parameters of the MA). This means in the MA case that beyond two lags the earnings growth is no longer autocorrelated, whereas the autocorrelation never disappears (it decreases geometrically with the lag length) in the AR context.

The distribution of the unobserved component $\ln \omega$ is drawn from a normal distribution with variance σ_ω equal to 0.7 throughout. κ in equation (8), which can be understood as the mean log earnings in the absence of taxation, is set equal to 12. These values are chosen to reproduce broadly the feature of the Swedish distribution of earnings in 2007, the year for which Ericson et al., (2015) describe the components of earnings, hours, and wages in Sweden in much detail.

We allow the share of the variance of the unobserved component that can be attributed to the permanent part, ϕ , to vary between 0.25 and 0.75 in increments of 0.25. In the extreme, when the share ϕ is equal to 1, individuals have exactly identical tastes in each time period and therefore their past choices are perfectly correlated with current choices. Lower levels of this share describe individuals with tastes that vary substantially from one period to the next. This determines whether the past quantities (earnings levels, net-of-tax rates, etc.) are good instruments for current net-of-tax rates.

Tax environment. In each year, we use a tax system that features two different marginal tax rates and a single kink point. Given these features, we simulate two different sets of tax environments, depending on whether or not the analyzed estimator uses the panel dimension of the data. Both the IV regression and the Indirect Inference estimators use the panel dimension, while the bunching estimator uses only cross-sectional information.

For the panel estimators, we choose four different tax environments that differ both in terms of tax rates and their time paths. We report the details of these tax environments in Table 17 in Appendix A. Our first tax environment (SW) is a simplified version of the Swedish income tax system over the period 2002-2013. In general, the Swedish income tax system consists of a local tax rate that applies to all income, and a national tax rate that applies to all income that exceeds a certain threshold income. We use the sum of the municipal and county level tax rates in one single Swedish municipality (Umeå) as the local tax rate and then add the national marginal tax rate. The kink point represents the income level above which individuals pay the national income tax, and varies over time. This first tax environment is a stylized version of the one a researcher would face in practice. Yet, it contains very little variation, as the tax rates only change once in 2005 (the kink points change every year).

In order to assess the role of the tax environment, our second environment (DK) uses a more variable sequence of tax rates. In particular, we use the lower two tax rates of the Danish tax system described in Kleven and Schultz, (2014), which is praised by the authors for its variability, in particular regarding the direction of tax rate changes. In order to keep the mean tax rates comparable to our first tax environment, we adjust the Danish tax rates accordingly. Effectively, we use a hybrid tax environment with the Danish variability of tax rates, but with the Swedish mean tax rates. As the direction of the tax rate change is considered crucial for the bias by some authors (see for example Weber, 2014), our third and fourth tax environment are sorted versions of the second environment. In the third environment, the spread between the two tax rates increases monotonously (DKi), while it decreases in the fourth environment (DKd).

For the bunching estimator, we do not vary the structure of the unobserved component. Instead, we apply two different tax environments: a large-kink (20 percentage points) and a small-kink (10 percentage points) environment, in which we keep the two tax rates constant over all years, so that only the location of the kink point changes. In all other aspects, the details of the simulations are left unchanged.

4. ESTIMATORS

We will now analyze the ability of the different methods used in the literature to measure α . In order to achieve that aim, we implement the methods that prevail in the literature on our simulated data. Furthermore, we implement our newly suggested Indirect Inference estimator and compare its performance to the prevalent methods.

Regression-Based Estimators (differences). We base our simulations of panel regression estimations on two influential studies, Gruber and Saez, (2002) (GS) and Weber, (2014) (Weber). Both studies use the same specification for their estimating equation, but differ in the instrumentation of the endogenous variables.

The core idea of the panel regression approaches is to regress the change in the tax units' incomes on the change of their net-of-tax rates between two periods of time, which we will refer to as start year and end year. The difference between start year and end year is typically between one and three years, and will be denoted by d . The literature has identified two core challenges to the regression approach, which are the endogeneity of the net-of-tax rate, and the correlation between the error term and start-year income (Auten and Carroll, 1999; Gruber and Saez, 2002; Saez, 1999, 2003; Weber, 2014).

GS was not the first, but the most influential study to address both challenges. In order to control for the endogeneity of the net-of-tax rate, they instrument the net-of-tax rate with a hypothetical net-of-tax rate based on tax units' base-year incomes. This base-year instrumentation was very influential and has been refined, but not substantially changed, by the subsequent literature. To control for the correlation between the error term and start-year income, GS control for

base-year income both in logs and as a spline in levels. They argue that two income controls are necessary because the error term can be correlated with the start-year income for two reasons: mean reversion and differential income trends across the income distribution. In addition, GS allow for income effects and include year and marital status dummies.

GS base their measurement of the ETI on an extended regression model specified as follows:

$$\begin{aligned} \Delta_d \ln wh_{it} = & \beta_0 + \alpha \Delta_d \ln \tau_{it}^c + \beta_1 \Delta_d \ln (wh_{it} - T_{it}) \\ & + \beta_2 \ln wh_{i,t-d} + \sum_{j=1}^{10} \delta_j S_j(wh_{i,t-d}) + \Delta_d \epsilon_{it} \end{aligned} \quad (10)$$

where β_0 is a constant (in practice, the constant may be year specific or account for differences in marital status), wh_{it} stands for individual i 's observed taxable income in year t , $T_{it} \equiv T_t(wh_{it}, \mathbf{x}_{it})$ describes the tax due in period t by individual i , and $\tau_{it}^c \equiv \tau_t^c[wh_{it}, \mathbf{x}_{it}]$ is individual i 's marginal net-of-tax rate in period t . Therefore $wh_{it} - T_{it}$ captures individual i 's net income and the behavioral reaction to its change corresponds to the income effect. Equation (10) contains two income controls, log start-year taxable income, $\ln wh_{i,t-d}$, and an income spline of the start-year taxable income, $\sum_{j=1}^{10} \delta_j S_j(wh_{i,t-d})$. GS argue that these two terms can control for correlations between the error term $\Delta_d \epsilon_{it}$ and start-year income $wh_{i,t-d}$. Note that this estimation equation differs from the model in Equation (8), since it now contains lagged gross income both in logs and as spline, and a measure of net income.

GS recognize that the unobserved component and the change of the net-of-tax rate are likely to be correlated so that the classical regression model assumption that $\mathbb{E}[\Delta_d \epsilon_{it} | \Delta_d \ln \tau_{it}^c, \ln wh_{i,t-d}] = 0$ does not hold in general. Hence, they instrument the change of the net-of-tax rate between period t and period $t - d$ with the hypothetical change of the net-of-tax rate keeping earnings fixed at the level observed in period $t - d$, $wh_{i,t-d}$. This hypothetical change $\Delta_d \ln \tilde{\tau}_t^c$ is defined as

$$\Delta_d \ln \tilde{\tau}_t^c \equiv \ln \tau_t^c[wh_{i,t-d}, \mathbf{x}_{it}] - \ln \tau_{t-d}^c[wh_{i,t-d}, \mathbf{x}_{i,t-d}],$$

and describes the marginal change to the tax system over time all else constant. By using $\Delta_d \ln \tilde{\tau}_t^c$ to explain the observed change of the net-of-tax rate, GS assume that $\mathbb{E}[\Delta_d \epsilon_{it} | \Delta_d \ln \tilde{\tau}_{it}^c] = 0$. Obviously the quality of this hypothetical change as an instrument remains an empirical question in practice.

In our setting, we simulate the GS estimator adjusted to the complexity of the model that generates our data, i.e., we do not control for factors that do not play a role in our simulated setting (such as marital status, income effects, and year effects). In addition, we do not include splines (our simulation experiment focuses on variations of the tax system over time as the only sources of variation for the observed distribution of earnings). We control for start-year log earnings in order to address mean reversion due to shocks on wage and the disutility of work. Our

application of the GS methodology to the simulated data is thus based on the model¹⁰

$$\Delta_d \ln wh_{it} = \beta_0 + \alpha \Delta_d \ln \tau_{it}^c + \beta \ln wh_{i,t-d} + \Delta_d \epsilon_{it}, \quad (11)$$

instrumenting the true change in the net-of-tax rate $\Delta_d \ln \tau_{it}^c$ with the hypothetical change $\Delta_d \ln \tilde{\tau}_t^c$.

Weber Instrumentation. Weber, (2014) argues that the instrument $\Delta_d \ln \tilde{\tau}_t^c$ is not exogenous as it relies on base-year income, which may be correlated with the unobserved term $\Delta_d \epsilon_{it}$.

Starting with the case where there is no transitory serial correlation, i.e., $\rho = 0$ or $\theta = 0$, Weber argues that functions of the start-year income $wh_{i,t-d}$ are endogenous as they are correlated with the start-year income. In particular, the GS instrument $\Delta_d \ln \tilde{\tau}_{i,t}^c$ is only exogenous if $\text{Cov}[\Delta_d \ln \tilde{\tau}_{i,t}^c, \Delta_d \epsilon_{it}] = 0$. If the predicted change in the net-of-tax rate depends monotonously on the start year's income level $wh_{i,t-d}$ then it is correlated with the start year's transitory income component. If the tax reform is a tax rate increase, all else equal, an increase in transitory income in the start year causes both an increase in the predicted tax rate change and a decrease in the transitory income change (the error component in first differences). The GS instrument is therefore negatively correlated with the error term and the estimate is lower than the true ETI. In the AR(1) case, the covariance between $\Delta_d \epsilon_{it}$ and $\ln wh_{i,t-d}$ decreases with the serial correlation of the transitory component. Therefore for moderate serial correlations any quantity defined using $\ln wh_{i,t-d}$ will covary with $\Delta_d \epsilon_{it}$ and the measurement of the ETI will be inconsistent. More generally, the direction of the IV estimation bias of the ETI depends on (i) the covariance between the net-of-tax rate change and the start year income level, and (ii) the direction of the tax rate change.

Weber argues that the following alternative instrument involving a deeper lag is preferable:

$$\Delta_{d,k} \ln \tilde{\tau}_{it}^c \equiv \ln \tau_t^c[wh_{i,t-k-d}, \mathbf{x}_{it}] - \ln \tau_{t-d}^c[wh_{i,t-k-d}, \mathbf{x}_{i,t-d}]. \quad (12)$$

This is the predicted change in the net-of-tax rate constructed using the tax unit's income k periods before the start year $t - d$. The GS instrument is obtained when $k = 0$. Weber shows that instruments constructed using earnings one period before the start year (i.e. such that $k = 1$) yield a consistent estimator if the unobserved component is not autocorrelated at any order. If the unobserved component follows a MA(i), Weber argues that instruments relying on earnings lagged $i+1$ or more periods yield consistent estimators of the ETI. Earnings lagged two or three periods satisfy this requirement in the case of our MA(1) process. For an AR process, even earnings observed in a distant past fail to satisfy the Weber

¹⁰Note that due to mean reversion, lagged income plays a role even if the data generating process is based on a quasi-linear utility function and shocks to wage and preferences are not serially correlated. In a more general model, this could make income effects hard to identify.

condition. Yet, in the case of an AR(1), Weber suggests that considering earnings in a distant enough past would provide "practically" acceptable instruments.¹¹

Using the same data, GS and Weber analyze a reform where the tax increases and the tax rate change increases with the income level. Weber's results suggest that her estimates of the ETI, keeping the specification constant, can be much larger if further income lags are used to construct the instrument. Yet, direct comparisons with GS are difficult as the estimates depend on details of the specifications (such as the selection of years, and whether or not an income spline is included as a control variable), and generally the variability around the estimates is large.

Controlling for Mean Reversion. It is interesting to note that the IV regression based approaches reviewed above simultaneously use lagged income as part of the construction of the instrumental variables and to control for mean reversion, i.e., the correlation between the residual in first differences and the start-year income level. In the ETI literature, mean reversion was first discussed by Carroll, 1998. In our view, while mean reversion may be the cause of some bias it is not clear that controlling for lagged income provides a solution. In fact, going back to our model in Equation (8), replacing $\ln \omega$ with $\sigma^P u_i^P + \sigma^T u_{it}^T$, we have

$$\begin{aligned}\ln wh_{it}^* &= \kappa + \alpha \ln \tau_{it}^c + \sigma^P u_i^P + \sigma^T u_{it}^T, \\ \ln wh_{it-d}^* &= \kappa + \alpha \ln \tau_{it-d}^c + \sigma^P u_i^P + \sigma^T u_{it-d}^T.\end{aligned}\tag{13}$$

which in first differences yields

$$\Delta_d \ln wh_{it}^* = \alpha \Delta_d \ln \tau_{it}^c + \sigma^T \Delta_d u_{it}^T.\tag{14}$$

As both $\Delta_d \ln wh_{it}^*$ and $\sigma^T \Delta_d u_{it}^T$ contain u_{it-d}^T , they are correlated (see also Kopcuk, 2003, 2005). This is the reason why GS add $\ln wh_{it-d}^*$ as a control variable. The question here is whether we expect adding $\ln wh_{it-d}^*$ as a control to reduce the bias. One easy way to look at the problem is to replace u_{it-d}^T in Equation (14) by using Equation (13):

$$\Delta_d \ln wh_{it}^* = \alpha \Delta_d \ln \tau_{it}^c + \sigma^T u_{it}^T - (\ln wh_{it-d}^* - (\kappa + \alpha \ln \tau_{it-d}^c + \sigma^P u_i^P)).\tag{15}$$

Hence, while Equation (15) reformulates the model such that it contains the base-year income instead of the base-year transitory income component in the spirit of GS, it contains an income component which cannot be observed (and which GS can not control for): the individual-specific effect u_i^P . The instrumentation strategy involving deeper lags of income to generate the instrument $\Delta_d \ln \tilde{\tau}_{i,t}^c$ in the context of Equation (14) is no longer adapted to construct a consistent estimator of α if

¹¹Weber uses an AR(j) process in her model setup, but argues that an AR and an MA process would be empirically indistinguishable. Instead, she tests empirically whether $\mathbb{E}[\ln y_{t-d-k} \Delta \ln \epsilon_t] = 0$ can be rejected.

the model is specified as Equation (15). This is so because the deeper income lags remain correlated with u_i^P and therefore $\Delta_d \ln \tilde{\tau}_{i,t}^c$ would be correlated with u_i^P . It is not clear that the latter estimates would be less biased than the first.

Bunching Estimators (levels). We assess the performance of two different bunching estimators. The first is the original estimator proposed by Saez, (2010). The second estimator is a parametric version based on the Saez framework, which assumes log-normality of the unobserved component. We give a more detailed description of both methods in a related paper (Aronsson et al., 2018).

Following Saez' intuition for the bunching estimator, imagine that the optimal level of earnings wh is an increasing function of an unobserved earnings component ω (which contains the wage and the disutility of work) as well as an increasing function of the net-of-tax rate (as defined in our model presented in Section 2). In a tax environment with two net-of-tax rates and one kink, define k as the earnings level at the kink, τ_1^c as the net-of-tax rate for earnings below the kink and τ_2^c as the (lower) net-of-tax rate for earnings above the kink. The earnings distributions below and above the kink will now follow different distributions: $wh(\tau_1^c, \omega)$ below the kink and $wh(\tau_2^c, \omega)$ above the kink. Note that the difference between the two distributions defines the ETI, as it describes the change in earnings due to a change in the net-of-tax rate. Further observe that for a particular value $\hat{\omega}$, $wh(\tau_1^c, \hat{\omega})$ exceeds $wh(\tau_2^c, \hat{\omega})$. Consequently, the value of ω for which $wh(\tau_1^c, \omega) = k$ is lower than the value of ω for which $wh(\tau_2^c, \omega) = k$. This defines a range of ω where the individual would choose an income above k under τ_1^c , but an income below k under τ_2^c . Let us define $\underline{\omega}$ as the lowest value and $\bar{\omega}$ as the highest value in that range. Individuals with an unobserved component $\omega \in [\underline{\omega}, \bar{\omega}]$ will, according to Saez, (2010), be observed at the kink.

Given this reasoning, it is possible to express the percentage B of tax units that bunch at the kink in the following way:

$$B = \int_k^{wh(\tau_1^c, \bar{\omega})} \tilde{f}(z; \tau_1^c) dz, \quad (16)$$

where $\tilde{f}(z; \tau^c)$ describes the earnings density given the net-of-tax rate τ^c . Saez approximates B using a trapezoidal approximation,

$$B \approx (wh(\tau_1^c, \bar{\omega}) - k) \left(\tilde{f}(k; \tau_2^c) \gamma(\tau_1^c, \tau_2^c, k) + \tilde{f}(k; \tau_1^c) \right) \frac{1}{2}. \quad (17)$$

For the isoelastic model (used by Saez, 2010, and described in part 2 in this paper) the optimal earnings level takes the form

$$wh(\tau, \omega) = \tau^{c\alpha} \omega.$$

Then, $wh(\tau_1^c, wh^{-1}(\tau_2^c, k)) - k = \left(\left(\frac{\tau_1^c}{\tau_2^c} \right)^\alpha - 1 \right) k$, $wh_z^{-1}(\tau^c, z) = \frac{1}{\tau^{c\alpha}}$, and $\gamma(\tau_1^c, \tau_2^c, k) = \left(\frac{\tau_1^c}{\tau_2^c} \right)^{-\alpha}$. Substitution in the approximate expression (17) gives Saez's equation (5)

exactly, which implicitly defines the ETI, α , given the observable parameters B , $\tilde{f}(k; \tau_1^c)$, and $\tilde{f}(k; \tau_2^c)\gamma(\tau_1^c, \tau_2^c, k)$. In practice, in order to measure these parameters, the researcher has to choose three intervals: one interval around the kink to define which tax units are considered to be part of B , and two symmetric intervals on each side of the kink interval that define the area based on which the densities $\tilde{f}(k; \tau_1^c)$ and $\tilde{f}(k; \tau_2^c)\gamma(\tau_1^c, \tau_2^c, k)$ are computed. When applying the Saez bunching estimator, we choose each of these intervals to be one percent of the kink income.

The parametric version relies on the log-normality of the distribution of the unobserved component. It is then straightforward to derive the likelihood of observation in the neighborhood of the kink and obtain maximum likelihood estimators of the parameters of the model, in particular the two density functions $\tilde{f}(k; \tau_1^c)$ and $\tilde{f}(k; \tau_2^c)$. Two possible estimators of the ETI can be directly derived from the maximum likelihood estimates. The first estimator corresponds exactly to Saez's estimator in the log normal case evaluated at the ML estimates

$$\hat{\alpha}_{SaezN} = \frac{2}{\hat{s} \ln \tau_1^c - \ln \tau_2^c} \frac{1}{\phi(\hat{s}(\ln k - \delta) - \hat{\lambda}_1) + \phi(\hat{s} \ln k - \hat{\lambda}_2)} \frac{\Phi(\hat{s} \ln k - \hat{\lambda}_2) - \Phi(\hat{s}(\ln k - \delta) - \hat{\lambda}_1)}{\phi(\hat{s}(\ln k - \delta) - \hat{\lambda}_1) + \phi(\hat{s} \ln k - \hat{\lambda}_2)}, \quad (18)$$

where \hat{s} , $\hat{\lambda}_1$, and $\hat{\lambda}_2$ estimate locations and scale parameters of the log-normal model ($\hat{\lambda}_1 > \hat{\lambda}_2$). The second estimator uses the behavioral model framework to obtain an estimator of the ETI as a function of the parameters directly

$$\hat{\alpha}_{norm} = \frac{\hat{\lambda}_1 - \hat{\lambda}_2}{\ln(\tau_1^c) - \ln(\tau_2^c)} \frac{1}{\hat{s}}. \quad (19)$$

We can show that these two ML based estimators are in strict relation with one another. In our simulations, we will use Saez' original estimator as in Equation (17), and the parametric version presented in Equation (19).

Indirect Inference (levels and differences). Given the specification of the earnings function (and in the absence of any information on the wage), all observations and all earnings histories are informative about the ETI. Yet, each of the conventional approaches described above uses only a fraction of the available information. The regression approach only focuses on differences and disregards the information on bunching, while the bunching approach only uses the cross-sectional information of observations near the kink point. Indirect inference allows us to combine the information from the earnings levels and earnings growth in order to obtain better behaved estimators of the ETI. While our particular focus here is on a simple behavioral model, the same principle can be extended to more demanding environments.

The indirect inference approach relies on an assumption concerning the data generating process. In our case, the data is generated according to the economic model structure (preferences and constraints) and the assumptions concerning the structure and distribution of the unobserved components. The maintained hypothesis throughout is that the modeling structure is the correct one, however the

exact parameter values which generate the sample data at hand are unknown. Denote Ξ the vector of parameters of the model (under the restrictions we discussed in the model section)

$$\Xi = (\alpha, \kappa, \sigma_\omega, \phi, \rho),$$

and denote Ξ_0 the particular parameter vector which generates the data. Ξ_0 is unknown and the statistical problem is concerned with its measurement.

On the basis of the observed data we can measure the auxiliary statistics, which contain information about Ξ_0 , and we denote this measurement $\hat{s}_0 \equiv \hat{s}(\Xi_0)$. Unfortunately, we are unable in general to retrieve directly an estimate of Ξ_0 from \hat{s}_0 . The relationship between the parameter vector which generate the data and the vector of auxiliary statistics is typically too complex or even beyond our ability to characterize completely to be able to "inverse" it.

The method of indirect inference suggests that we can obtain good estimates of Ξ_0 by using the modeling structure to simulate synthetic data and calculate the vector of auxiliary statistics on the synthetic data given a guess, Ξ_g , for the value of the parameter vector that generates the data. This simulation step provides us with a simulated measurement $\hat{s}(\Xi_g)$ of the auxiliary statistics.

While it is not feasible to determine directly how far a given guess Ξ_g is from the true value of the parameters Ξ_0 , we are able to evaluate the distance between the vectors of auxiliary statistics \hat{s}_0 and $\hat{s}(\Xi_g)$. The intuition is therefore that if Ξ_g is such that $\hat{s}(\Xi_g) \approx \hat{s}_0$, i.e., Ξ_g allows the simulated auxiliary statistics to match the observed values of the auxiliary statistics, then $\Xi_g \approx \Xi_0$. Ξ_g is then a good estimate of the true value of the parameter vector.

An indirect inference (I-I) estimator for Ξ_0 is then the vector $\hat{\Xi}_{\text{I-I}}(\hat{s}_0)$ which minimizes the distance $Q(\Xi_g, \hat{s}_0, A)$ between the observed and the simulated auxiliary statistics. A natural distance is the Euclidian distance

$$Q(\Xi_g, \hat{s}_0, A) = (\hat{s}_0 - \hat{s}(\Xi_g))' A (\hat{s}_0 - \hat{s}(\Xi_g)),$$

where A is a constant positive definite matrix. In effect, we are using the minimization of the objective $Q(\Xi_g, \hat{s}_0, A)$ to solve approximately the equation $\hat{s}_0 = \hat{s}(\Xi)$ for Ξ .

The large sample theory in terms of the number of observations and the number of simulated observations is well developed and will apply to our context directly. In particular, the I-I estimator is asymptotically normally distributed for any positive definite A . Given a sensible choice for A , estimates of the asymptotic variance covariance matrix of the estimator exist and can be calculated from quantities that are derived from the objective function at its optimum. Gouriéroux et al., (1993) or Jiang and Turnbull, (2004) provide presentations of the large sample theoretical properties of the method and describe a wide array of applications.

In the current paper we focus on showing that I-I can provide a feasible solution to the issue of unbiased estimation. We will not pursue the question of producing the most efficient I-I estimator. We therefore limit ourselves to the case where A is set to the identity matrix. Hence any improvements that are obtained in terms

of bias by our implementation of I-I do not preclude further potential efficiency gains that can be obtained by a more adapted (but more demanding) choice of A .

In practice we use the default GMM optimizer in `Stata` to compute the I-I estimates. The starting values are obtained systematically from the auxiliary statistics of the sample, such that all starting values are endogenously determined based on the data.

Choice of auxiliary statistics. We consider auxiliary statistics from both types of conventional estimators, i.e., cross sectional statistics including bunching behavior, and panel statistics such as the autocorrelation coefficient of taxable income. Furthermore, we focus on statistics which can be calculated directly from the sample data (and therefore from the simulated data). Our choice of auxiliary statistics is guided mainly by the principle of analogy. Since we estimate all the parameters of the model, i.e., the ETI as well as all other parameters which arise from our distributional assumptions (location and scale), the composition of the unobserved component as well as the dynamics of the transitory component, we require auxiliary statistics which will capture separately distinct features of the observed distribution. To achieve this, we will use five sets of auxiliary statistics: means of income levels and growth, the proportions of tax units above and below the kink point, transitions between the two tax brackets, as well as variances and covariances of income levels, income growth, and the net-of-tax rate. We describe these statistics in detail in appendix C.

5. RESULTS

Overall, our simulation results show that the conventional estimators can be very variable, and in the case of IV-regression estimations substantially biased. We first discuss the regression-based estimators, then turn to the bunching estimators, and finally show the results for the indirect inference estimator.

Regression-based estimators. The regression-based estimators we consider are based on Equation (11) estimated using instrumental variables procedures. The performance differences between estimators arise from the specific construction of the set of instruments for the change in the net-of-tax rate and base-year income control. All instruments are constructed using lagged income/earnings. Distinct estimators are obtained as the number of lags prior to base-year increases. A lag of length zero corresponds to the set of instruments Gruber and Saez (2002) (GS) propose, while higher lags correspond to various forms of Weber’s (2014) instruments. In the following, we first consider the simulated distribution of the IV estimator using the GS instrumentation (with lag length $k = 0$). We increase the lag length to consider the effect of the Weber instrumentation with lag length $k = 3$.

Figure 2 illustrates results for the GS estimator in four different tax environments as described in detail in section 3. In all histograms presented here, the true ETI in the simulation is 0.6 (marked by a vertical line), and 50% of the

variance of the unobserved component is within individuals.¹² We find a large downward bias in all tax environments. The average relative bias is 69% for the most variable tax environment (DK), and increases to 86% in the least variable tax environment (SW). Overall, independently of the tax environments the bias is large. This result carries through for other parameter choices (see Tables 1 and 5). These results suggest that the base-year instrumentation used by GS produces substantially biased estimates, even though results across tax environments may be similar.

We now turn to the effect of the instrument choice on the distribution of the estimated ETI parameter. Figure 3 illustrates how the distribution of the estimator changes as we increase the lag lengths of the instrument for the net-of-tax rate and base-year income. The tax environment in all histograms in Figure 3 is the most variable tax environment (DK), i.e., the environment most likely to yield precise estimates. The top left panel shows results for a lag length of zero, which corresponds to the GS estimator (the top left panel in Figure 2). The top right panel presents the distribution of the IV estimator when the earnings are lagged once to construct the instrument, while the bottom panels show histograms of the estimates for earnings lagged two and three years before the base year.

We find that with increasing lag length, the bias decreases considerably. For the MA(1) error structure, the median of the estimated ETI parameters deviates 7% from the true ETI parameter for $k = 1$, while it is practically unbiased for $k > 1$. These results confirm the argument put forward by Weber that using pre-base-year lagged income for instrumentation reduces the bias of the estimator. Yet, the decreasing bias is accompanied by a dramatic increase in the variability of the estimator. Based on the instruments discussed above, the possible values of the ETI range roughly between zero and one, which corresponds to the range of parameter sizes on the left part of the Laffer curve (i.e., to the left of and on the top). Therefore, the Weber instrumentation applied to a dataset of comparable size to ours (10k individuals observed for 12 years) is not guaranteed to provide a precisely estimated value; instead, there is considerable uncertainty around the true value of the parameter. The values of the standard deviation and of the interquartile range of the simulated distribution reported in tables 2 to 4 for the AR(1) case and in table 6 to 8 for the MA(1) case show that the variability of the estimator is substantial in all tax environments. These results are also consistent with the large standard errors reported by Weber using the data analyzed by GS.¹³

¹²The results for other parameter choices (ETI values, variance compositions, and autocorrelation structures) are similar and are reported in Tables 1 to 8.

¹³Weber's data comprises roughly 6000 annual observations, and standard errors range around 0.3. We report observation numbers and standard errors of several IV regression studies in Table 18 in Appendix B. The reported estimates show that the variability of the estimators produced by our simulations, adjusted for sample size, is of comparable size to the range of observed empirical results.

Figures 4 and 5 show histograms for the Weber estimator when $k = 3$, for distinct tax environment and autocorrelation structure of the unobserved component. We can make three clear observations. First, the precision of the estimator depends on the tax environment. The bottom right panel of each figure shows results for the stylized Swedish tax environment, which has little tax variability, while the top left panel shows results for a hybrid tax environment with Danish variability, normalized to Swedish mean tax rates. The top right and bottom left panel show results for ordered versions of the hybrid tax environment, where the difference between the tax rates is either monotonously increasing (top right) or decreasing (bottom left). With both AR(1) and MA(1) error structures, the hybrid (and most variable) tax environment performs best in terms of precision, while the results for the least variable tax environment indicate a severe bias, with a relative bias exceeding 100%, and an interquartile range that exceeds the parameter size by an order of magnitude.

Second, the statistics show that for the three variable tax environments, the estimator is considerably less biased in the MA(1) case than in the AR(1) case. Both median and average estimates in the MA(1) case are within a close range (at most 6% deviation) of the correct value for the first three tax environments (even though unpredictable in the fourth), while the relative bias in the AR(1) case varies roughly between 5% and 10%. These differences are expected as the instruments are truly exogenous in the MA(1) case but not in the AR(1) case.

Third, even in the most variable tax environment estimates range roughly from zero to one. The interquartile range of the estimates is roughly of the same order of magnitude as the parameter value, putting serious doubt on the suitability of the estimator. Thus, the results for the regression estimators show that the endogenous instrument of the GS estimator leads to substantial bias, while lagged income instruments substantially decrease the precision. The imprecise nature of the estimates is consistent with the lack of robustness of the estimates obtained in the literature.

Bunching estimator. Turning to the two bunching estimators (Saez and lognormal, as described in part 4 and Aronsson et al., 2018), Figure 6 shows histograms for a "true" ETI value of 0.6 for selected years. The corresponding statistics for all years are reported in tables 9 to 11. As the bunching estimators are cross-sectional, we adjust our tax environments, as neither the autocorrelation process of the unobserved component nor the time pattern of tax rates is relevant. Instead, the tax environments we analyze here feature a larger and a smaller kink. In both panels on the left-hand side, the tax environment features a small change in the net-of-tax rate at the kink (10 percentage points), while on the right-hand side the change is large (20 percentage points). From top to bottom, the tax year increases, which in our data means that the kink point corresponds to a higher income level, with the wage distribution unchanged.

We obtain two main results: First, the bias of both estimators is small. The ML estimator based on the truncated log-normal distribution performs slightly better

in terms of bias in general. Notably, the estimators do not consistently perform better in the large-kink environment. This result is not entirely unexpected, as the observation numbers to the right of the kink are lower for the larger kink (given the size of the measurement intervals around the kink), adding noise to the estimation of the density on the right-hand-side of the kink. Second, both estimators are considerably more precise than the Weber estimators, and the ML estimator performs better than the original Saez estimator. The ML estimator's standard deviation and inter-quartile range are roughly an order of magnitude smaller than the same statistics for the Weber estimators.

Indirect Inference estimator. Turning to the I-I estimator, Figure 7 shows histograms for a true ETI of 0.6 and an MA(1) error process. Statistics for all specifications are reported in Tables (12) and (13). Compared to the conventional estimators, the I-I estimator is almost unbiased and very precise. The average relative bias is around 1% of the true parameter value, and both the standard deviation and the inter-quartile range are less than half the values of the log-normal bunching estimator. These results are independent of the tax environment. Furthermore, the results are similar for the AR(1) error process, suggesting that the precision of the I-I estimator does not depend on the autocorrelation structure.

In our baseline specifications, we thus find that the IV estimator has the largest bias and is the least precise. It appears that the IV estimator can work well in terms of bias in very variable tax environments, but it requires a large number of observations to perform well in terms of precision. On the other hand, under the same conditions both the bunching estimator and the I-I estimator are practically unbiased, with the I-I estimator being the least variable.

Notably, our baseline simulations do not suggest that bunching-based estimates are smaller than regression-based estimates in general. However, as any regression estimate is based on one particular realization of individual errors, bunching estimates can of course potentially be smaller or larger than the regression-based IV estimates on any particular dataset, either because they are more precise, or because the populations who identify either estimate differ. In the following robustness analysis, we suggest that optimization frictions can explain a systematic difference that causes the IV estimates to exceed the bunching estimates.

Robustness. Let us discuss the robustness of the three estimators by considering two distinct modifications of the data generating process: one which modifies the distribution of the unobserved component, and the other where the behavioral model is modified.

The first scenario tests the sensitivity of the estimators to a departure from the log-normality assumption. This modification is motivated by the thought that both the log-normal bunching estimator and the I-I estimator use distributional information of the unobserved component. We are therefore interested in how the results change if the true distribution is not log-normal, while the estimators assume a log-normal distribution. More specifically, the modified innovations used

to generate the data are distributed according to a scaled log-student distribution, while the simulated data within the I-I estimation are assumed to be log-normally distributed. We allow the number of degrees of freedom of the distribution of the innovations to change in two ways. In a first case, we set it equal to 5 so that the distribution of the innovations exhibits fatter tails than the log-normal distribution while insuring that the first four moments of the distribution exist. This is the most extreme departure from log-normality that we consider. In a second intermediate case we set the number of degrees of freedom equal to 10 (the first nine moments of the distribution of the innovations exist in this case). In both cases, we scale the innovations such that their variance is equal to 1 and then proceed with the data generation as we did above.

In the second scenario, we are interested in the effect on the results if only a portion of the population reacts to taxation. Such a case can be motivated either by optimization frictions or by a share of true non-responders in the population. The data is generated such that only a fixed proportion of the population is assumed to be responsive to the tax system; meaning, in turn, that a fixed proportion of the population is non-responsive to changes in marginal incentives. We set the proportion of the population that responds to the tax system to either 50 per cent or 75 per cent (while maintaining the log-normality of the innovations). This property is individual specific and not time dependent. Hence we do not generate data where individuals are alternatively responding and then not responding over their life cycle. Furthermore when simulated individuals do not respond to the tax system, we adjust the average earnings among non-respondents so that it matches that of respondents.

The consequences of these changes to the data generating process are analyzed in Tables (14) to (16), where the ETI is set to 0.6 and the share of the individual-specific effect in the unobserved component is set to 50 per cent.

We observe that the first modification (the departure from the log-normality assumption) leaves the results for the IV estimator and the bunching estimator essentially unchanged. For the I-I estimator, it slightly increases both the bias and the variability. The variability is now roughly a third (AR case) or a quarter (MA case) of the variability of the IV estimator, and does not exceed (but mostly undercut) the variability of the bunching estimator. Notably, this is the case for log-Student's distributions with both 10 and 5 degrees of freedom. The distributional misinformation of the I-I estimator reduces its precision, yet it still outperforms the other estimators.

The second modification (the introduction of non-responding tax units) has very different effects on the three estimators. It has little influence on the results of the IV estimator, albeit it slightly increases both the bias and the variability. By contrast, the estimates of the bunching estimator and the I-I estimator are scaled down roughly in proportion to the share of responders. That is, while the IV estimator measures the ETI of the responders only, both the bunching estimator and the I-I estimator measure the average ETI over all tax units. This

result can explain the differences the literature finds between the results of the bunching estimator and the results of the IV estimator. If there is a share of non-responders, for instance due to optimization frictions, the bunching method – and other methods which rely on the structure of the model to explain the data, like I-I – will on average generate smaller estimates.¹⁴

6. CONCLUSION

This paper departed from the observations that (i) the prevailing estimates of the ETI differ considerably across studies, and (ii) there are no earlier attempts of assessing the performance of the estimators used in the literature. We have used Monte Carlo simulations to assess the bias and precision of the prevalent methods of estimating the ETI: the regression method (based on different assumptions of instrumental variables) and the bunching method. We have also examined the performance of an alternative, simulation-based estimator relying on the method of indirect inference.

In order to simulate data generated by utility maximization, we assumed that labor constitutes the only income source and presented a structural labor supply model producing a log-linear labor supply function and thus also a log-linear income equation (where income is defined as the hourly wage rate times the hours of work). Albeit simple in structure, this model accords well with those used in the literature on the ETI. When simulating the data, we distinguished between different tax environments reflecting (a simplified version of) the Swedish tax system for the period 2002-2013 as well as three different mixtures of the Swedish and Danish tax systems (in order to increase the number of tax rate changes). We also allowed for both MA(1) and AR(1) processes for the unobserved component of the model.

To assess the regression-based estimators, we controlled for lagged income in order to address mean reversion as well as applied the instrumentation techniques suggested by Carroll, (1998), Gruber and Saez, (2002), and Weber, (2014), respectively. The latter produces an exogenous instrument for the net-of-tax-rate under more general conditions. We also assessed the performance of two bunching estimators; one based on Saez, 2010 and the other on our parametric model (as presented in part 4 as well as in Aronsson et al., 2018). The indirect inference estimator was derived by minimizing the (Euclidian) distance between the observed and simulated auxiliary statistics, where the auxiliary statistics are chosen to capture distinct features of the observed distribution. Our concern here was to examine whether an indirect inference estimator can provide a solution to the problem of unbiased estimation per se; not to try to find the most efficient estimator.

Finally, we modified our simulations in two dimensions to test the robustness of our results. We first changed the distribution of the unobserved component to

¹⁴We thank Etienne Lehmann for sharing this suspicion with us.

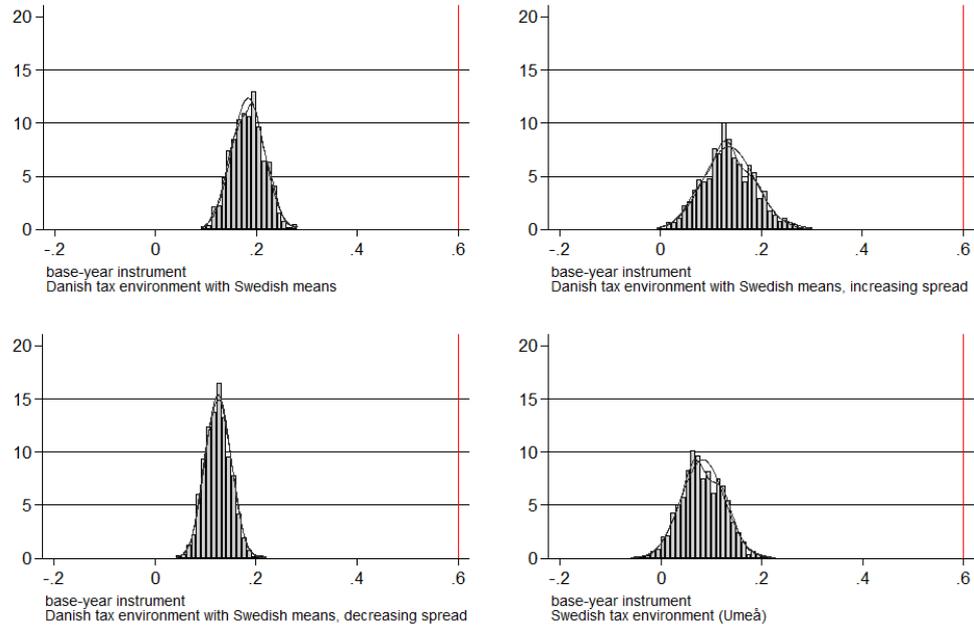
test the robustness of the indirect inference estimator with respect to distributional assumptions. We then assumed that only a fraction of the tax units respond to taxation, to test the performance of the estimators when there are non-responders, for example due to optimization friction.

The take home message of the simulations can be summarized by the following seven main results. First, the precision of the estimators based on the regression method largely depends on the underlying tax environment as well as on the exact process of the unobserved component. Albeit not very surprising in itself, this contributes to explain why estimates based on such approaches differ considerably between earlier studies. Second, the bias of these estimators largely depends on the process of the unobserved component. Focusing on the approach to instrumentation suggested by Weber, (2014), we found that the estimators are considerably more biased in the AR than the MA case, because the instrumentation procedure produces exogenous instruments only in the MA case. Third, our experimentation with the lag-lengths of the instruments for the net-of-tax-rate and base-year income, respectively, showed a possible trade-off between unbiasedness and precision. Whereas the Gruber-Saez estimator is quite precise and heavily biased (such that the correct parameter value is outside the estimated confidence interval), the Weber approach leads to considerably less bias combined with a large decrease in precision (compared to the Gruber-Saez estimator). We interpret the first three results to suggest that the regression-based methods do not in general perform well in our controlled environment.

Fourth, the bunching estimators performed quite well here, despite their property to be local in nature; their bias is relatively small and they are much more precise than the regression estimators based on Weber’s instrumentation approach. This result is reasonable, considering that the bunching estimators are cross-sectional, while the problems of bias and imprecision of the regression-based estimators are related to the time-series properties of the unobserved component and the development over time of the net-of-tax-rates. The local nature of the bunching estimator is not problematic in this specific case, as our simulated population is homogeneous with respect to the ETI. Fifth, our suggested indirect inference estimator performs very well in the model environment set out above. To be more specific, our results showed that this estimator is almost unbiased and very precise in each of the four tax environments. Sixth, a change in the true distributional form of the unobserved component moderately increases the bias and the variability of the indirect inference estimator; yet, given the data generating process, it is still more precise than the alternative estimators. Seventh, the results of the IV estimator remain essentially unchanged even if we introduce a share of non-respondents, while the bunching estimators and the indirect inference estimator measure a smaller ETI, which is scaled down roughly in proportion to the share of respondents. This result can explain the different magnitudes of ETI measures between the IV and the bunching estimators found in the literature.

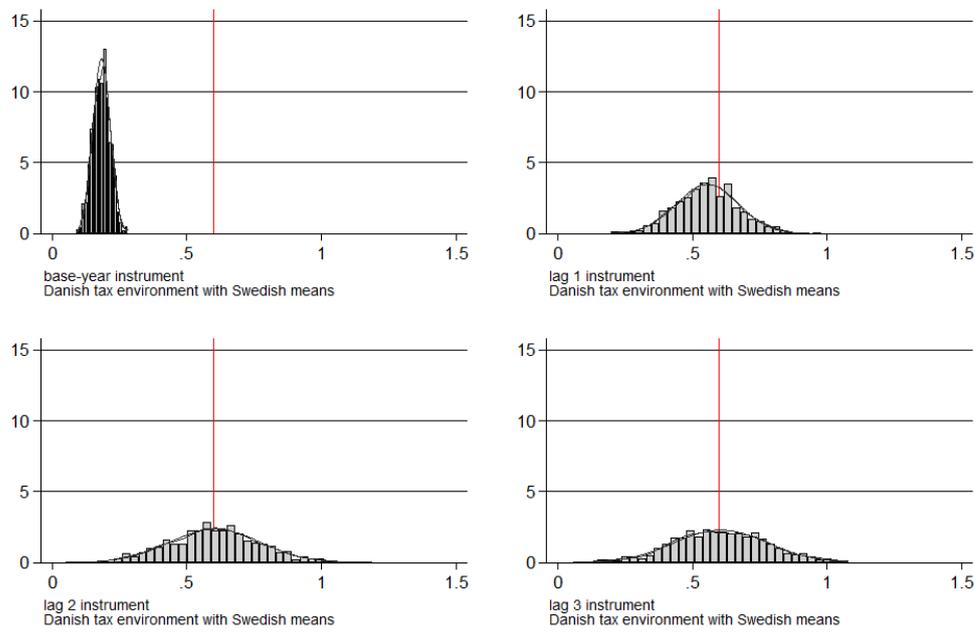
Future research may take several possible directions. Given the results presented above, one obvious direction would be to find the most efficient indirect inference estimator, while another would be to apply the indirect inference estimator to real-world data. The latter is interesting because the parametric model used above gives a taxable income equation similar to those used in much earlier research on the ETI, and the indirect inference method seems to perform very well in this particular environment. Yet, an even more interesting extension would be to evaluate how different estimators perform in more realistic model environments with more than one behavioral margin. The main argument for studying income formation - instead of the more narrow labor supply model - is to allow for several behavioral margins (in addition to work hours), such as wage formation, tax avoidance, and intertemporal choices. To arrive at unbiased estimates of the behavioral responses to taxation, both the underlying behavioral model and the estimation method (as well as the links between them) are important. We hope to address all these questions in the future.

FIGURE 2. Effect of tax environment GS estimator with base-year instrumentation ($k=0$) MA(1)



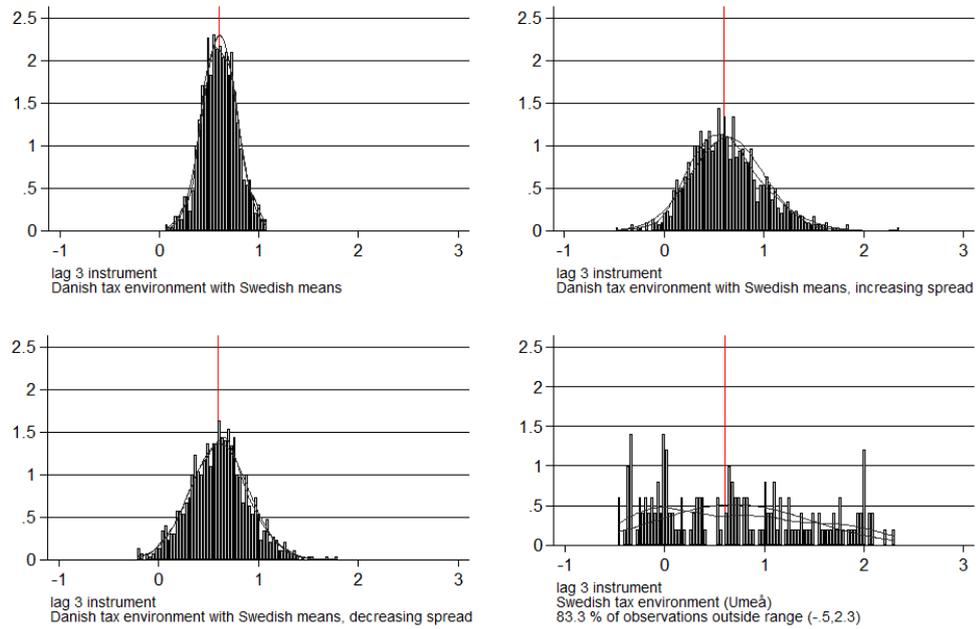
Note: Histograms of Gruber Saez estimator as given in Equation (11) with original start-year instrumentation which equals Weber instrumentation with lag length ($k = 0$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment given in the graph note. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See Table (17) and section (3) for details. See Table (5) for precise figures.

FIGURE 3. Effect of instrumentation lag in GS estimator with base-year and Weber instrumentation ($k=0,1,2,3$) MA(1)



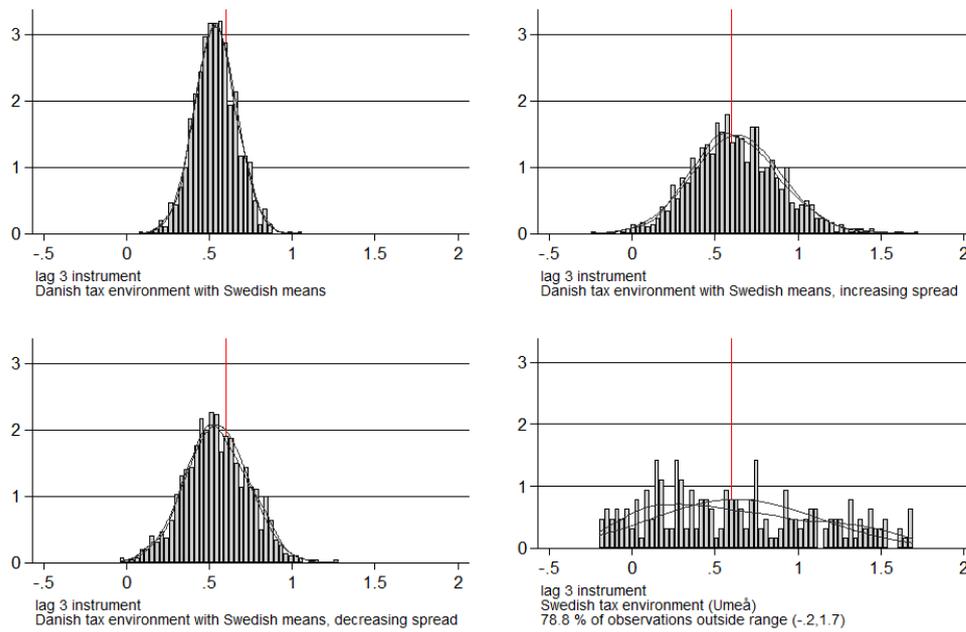
Note: Histograms of Gruber Saez estimator as given in Equation (11) with instrumentation lag length varying between ($k = 0, 1, 2, 3$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment given in the graph note. Tax environment: DK Danish tax variability with Swedish tax rate means. See Table (17) and section (3) for details. See Tables (5) to (8) for precise figures.

FIGURE 4. Effect of tax environment GS estimator with Weber instrumentation ($k=3$) MA(1)



Note: Histograms of Gruber Saez estimator as given in equation (11) with Weber instrumentation lag length ($k = 3$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See Table (17) and section (3) for details. See Table (8) for precise figures.

FIGURE 5. Effect of tax environment GS estimator with Weber instrumentation ($k=3$) AR(1)



Note: Histograms of Gruber Saez estimator as given in equation (11) with Weber instrumentation lag length ($k = 3$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See Table (17) and section (3) for details. See Table (1) for precise figures.

TABLE 1. Selected statistics of the simulated distribution of Gruber Saez estimator with base-year instrumentation ($k=0$), AR(1)

ETI	FE	TE	eti results				bias		relative bias	
			mean	sd	p50	iqr	mean	p50	mean	p50
30	25	DK	0.10	0.03	0.10	0.04	0.20	0.20	0.66	0.66
		DKi	0.07	0.06	0.07	0.08	0.23	0.23	0.77	0.77
		DKd	0.07	0.03	0.07	0.04	0.23	0.23	0.77	0.77
		SW	0.04	0.05	0.04	0.06	0.26	0.26	0.86	0.86
	50	DK	0.11	0.03	0.11	0.04	0.19	0.19	0.63	0.64
		DKi	0.07	0.05	0.07	0.07	0.23	0.23	0.76	0.76
		DKd	0.08	0.02	0.08	0.03	0.22	0.22	0.74	0.74
		SW	0.05	0.04	0.05	0.06	0.25	0.25	0.84	0.84
	75	DK	0.12	0.02	0.12	0.03	0.18	0.18	0.60	0.60
		DKi	0.08	0.04	0.07	0.06	0.22	0.23	0.74	0.75
		DKd	0.08	0.02	0.08	0.03	0.22	0.22	0.72	0.72
		SW	0.05	0.04	0.05	0.05	0.25	0.25	0.82	0.82
60	25	DK	0.18	0.03	0.18	0.04	0.42	0.42	0.70	0.70
		DKi	0.14	0.05	0.14	0.06	0.46	0.46	0.77	0.76
		DKd	0.11	0.03	0.11	0.03	0.49	0.49	0.81	0.81
		SW	0.08	0.04	0.08	0.05	0.52	0.52	0.86	0.87
	50	DK	0.19	0.03	0.19	0.04	0.41	0.41	0.69	0.69
		DKi	0.14	0.04	0.14	0.06	0.46	0.46	0.77	0.76
		DKd	0.12	0.02	0.12	0.03	0.48	0.48	0.80	0.80
		SW	0.08	0.04	0.08	0.05	0.52	0.52	0.86	0.86
	75	DK	0.20	0.02	0.20	0.03	0.40	0.40	0.67	0.67
		DKi	0.14	0.03	0.14	0.05	0.46	0.46	0.76	0.76
		DKd	0.13	0.02	0.13	0.02	0.47	0.47	0.79	0.79
		SW	0.09	0.03	0.09	0.04	0.51	0.51	0.85	0.86
90	25	DK	0.26	0.03	0.26	0.04	0.64	0.64	0.71	0.71
		DKi	0.22	0.04	0.22	0.06	0.68	0.68	0.76	0.76
		DKd	0.15	0.02	0.15	0.03	0.75	0.75	0.83	0.83
		SW	0.11	0.04	0.11	0.05	0.79	0.79	0.87	0.87
	50	DK	0.28	0.03	0.28	0.04	0.62	0.62	0.69	0.69
		DKi	0.22	0.04	0.22	0.05	0.68	0.68	0.76	0.76
		DKd	0.16	0.02	0.16	0.03	0.74	0.74	0.82	0.82
		SW	0.12	0.04	0.12	0.05	0.78	0.78	0.87	0.87
	75	DK	0.28	0.02	0.28	0.03	0.62	0.62	0.69	0.69
		DKi	0.22	0.03	0.22	0.04	0.68	0.68	0.76	0.76
		DKd	0.16	0.02	0.16	0.02	0.74	0.74	0.82	0.82
		SW	0.12	0.03	0.12	0.04	0.78	0.78	0.87	0.87

Note: Gruber Saez estimator as given in equation (11) with base-year instrumentation ($k = 0$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

TABLE 2. Selected statistics of the simulated distribution of Gruber Saez estimator with Weber instrumentation ($k=1$), AR(1)

ETI	FE	TE	eti results				bias		relative bias	
			mean	sd	p50	iqr	mean	p50	mean	p50
30	25	DK	0.28	0.11	0.28	0.15	0.02	0.02	0.05	0.07
		DKi	0.34	0.38	0.30	0.48	-0.04	0.00	-0.12	-0.02
		DKd	0.20	0.17	0.20	0.24	0.10	0.10	0.32	0.32
		SW	0.48	1.52	0.23	1.63	-0.18	0.07	-0.60	0.22
	50	DK	0.29	0.10	0.30	0.14	0.01	0.00	0.02	0.02
		DKi	0.32	0.34	0.29	0.44	-0.02	0.01	-0.07	0.03
		DKd	0.23	0.15	0.23	0.21	0.07	0.07	0.22	0.23
		SW	0.44	1.35	0.21	1.28	-0.14	0.09	-0.45	0.29
	75	DK	0.30	0.07	0.30	0.10	0.00	0.00	0.00	0.00
		DKi	0.30	0.24	0.28	0.32	0.00	0.02	-0.01	0.06
		DKd	0.27	0.11	0.26	0.14	0.03	0.04	0.11	0.13
		SW	0.30	0.79	0.16	0.78	0.00	0.14	-0.01	0.48
60	25	DK	0.56	0.10	0.56	0.14	0.04	0.04	0.06	0.07
		DKi	0.58	0.30	0.55	0.38	0.02	0.05	0.03	0.08
		DKd	0.40	0.17	0.39	0.23	0.20	0.21	0.33	0.34
		SW	0.46	1.16	0.35	1.14	0.14	0.25	0.23	0.42
	50	DK	0.57	0.10	0.58	0.12	0.03	0.02	0.04	0.04
		DKi	0.58	0.26	0.57	0.35	0.02	0.03	0.03	0.04
		DKd	0.45	0.16	0.45	0.20	0.15	0.15	0.25	0.25
		SW	0.49	1.47	0.34	0.99	0.11	0.26	0.18	0.43
	75	DK	0.59	0.07	0.59	0.10	0.01	0.01	0.02	0.02
		DKi	0.60	0.19	0.59	0.25	0.00	0.01	0.00	0.02
		DKd	0.52	0.12	0.52	0.16	0.08	0.08	0.13	0.13
		SW	0.54	0.65	0.45	0.73	0.06	0.15	0.11	0.25
90	25	DK	0.87	0.10	0.87	0.13	0.03	0.03	0.04	0.04
		DKi	0.82	0.23	0.80	0.30	0.08	0.10	0.09	0.11
		DKd	0.61	0.18	0.61	0.23	0.29	0.29	0.32	0.33
		SW	0.54	0.92	0.44	1.07	0.36	0.46	0.40	0.51
	50	DK	0.88	0.09	0.87	0.13	0.02	0.03	0.03	0.03
		DKi	0.84	0.21	0.83	0.28	0.06	0.07	0.07	0.07
		DKd	0.69	0.16	0.68	0.21	0.21	0.22	0.24	0.24
		SW	0.63	0.78	0.53	0.94	0.27	0.37	0.30	0.42
	75	DK	0.89	0.07	0.89	0.10	0.01	0.01	0.01	0.01
		DKi	0.87	0.17	0.86	0.22	0.03	0.04	0.03	0.05
		DKd	0.79	0.13	0.78	0.17	0.11	0.12	0.12	0.13
		SW	0.77	0.84	0.64	0.71	0.13	0.26	0.14	0.29

Note: Gruber Saez estimator as given in equation (11) with Weber instrumentation ($k = 1$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

TABLE 3. Selected statistics of the simulated distribution of Gruber Saez estimator with Weber instrumentation ($k=2$), AR(1)

ETI	FE	TE	eti results				bias		relative bias	
			mean	sd	p50	iqr	mean	p50	mean	p50
30	25	DK	0.27	0.14	0.27	0.20	0.03	0.03	0.09	0.09
		DKi	0.39	0.35	0.37	0.50	-0.09	-0.07	-0.29	-0.25
		DKd	0.25	0.25	0.25	0.33	0.05	0.05	0.15	0.18
		SW	0.56	11.27	0.21	2.41	-0.26	0.09	-0.88	0.31
	50	DK	0.28	0.13	0.28	0.19	0.02	0.02	0.07	0.08
		DKi	0.37	0.31	0.36	0.40	-0.07	-0.06	-0.22	-0.18
		DKd	0.26	0.22	0.25	0.29	0.04	0.05	0.15	0.16
		SW	-3.05	123.25	0.21	1.87	3.35	0.09	11.17	0.31
	75	DK	0.29	0.10	0.29	0.13	0.01	0.01	0.03	0.03
		DKi	0.34	0.24	0.32	0.31	-0.04	-0.02	-0.14	-0.07
		DKd	0.28	0.16	0.28	0.21	0.02	0.02	0.08	0.07
		SW	0.56	2.22	0.24	1.25	-0.26	0.06	-0.85	0.21
60	25	DK	0.54	0.12	0.53	0.16	0.06	0.07	0.11	0.12
		DKi	0.74	0.29	0.72	0.38	-0.14	-0.12	-0.23	-0.19
		DKd	0.50	0.21	0.49	0.28	0.10	0.11	0.17	0.18
		SW	0.88	7.22	0.34	2.04	-0.28	0.26	-0.47	0.43
	50	DK	0.56	0.12	0.56	0.16	0.04	0.04	0.07	0.06
		DKi	0.70	0.27	0.68	0.36	-0.10	-0.08	-0.16	-0.14
		DKd	0.53	0.19	0.52	0.25	0.07	0.08	0.12	0.13
		SW	0.96	9.41	0.49	1.64	-0.36	0.11	-0.60	0.19
	75	DK	0.58	0.10	0.58	0.13	0.02	0.02	0.04	0.04
		DKi	0.66	0.21	0.64	0.28	-0.06	-0.04	-0.10	-0.07
		DKd	0.56	0.14	0.55	0.19	0.04	0.05	0.07	0.08
		SW	0.81	2.38	0.47	1.13	-0.21	0.13	-0.35	0.21
90	25	DK	0.82	0.12	0.82	0.15	0.08	0.08	0.09	0.09
		DKi	1.06	0.23	1.05	0.30	-0.16	-0.15	-0.18	-0.17
		DKd	0.79	0.19	0.79	0.26	0.11	0.11	0.13	0.12
		SW	0.89	3.42	0.61	1.91	0.01	0.29	0.01	0.33
	50	DK	0.85	0.11	0.84	0.15	0.05	0.06	0.06	0.06
		DKi	1.01	0.22	1.00	0.30	-0.11	-0.10	-0.12	-0.11
		DKd	0.82	0.17	0.82	0.22	0.08	0.08	0.09	0.09
		SW	1.12	5.27	0.69	1.61	-0.22	0.21	-0.24	0.24
	75	DK	0.87	0.09	0.87	0.12	0.03	0.03	0.03	0.04
		DKi	0.97	0.18	0.96	0.24	-0.07	-0.06	-0.08	-0.07
		DKd	0.85	0.13	0.85	0.17	0.05	0.05	0.05	0.06
		SW	1.00	1.43	0.77	1.09	-0.10	0.13	-0.11	0.14

Note: Gruber Saez estimator as given in equation (11) with Weber instrumentation ($k = 2$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

TABLE 4. Selected statistics of the simulated distribution of Gruber Saez estimator with Weber instrumentation ($k=3$), AR(1)

ETI	FE	TE	eti results				bias		relative bias	
			mean	sd	p50	iqr	mean	p50	mean	p50
30	25	DK	0.23	0.15	0.23	0.20	0.07	0.07	0.22	0.23
		DKi	0.35	0.38	0.33	0.49	-0.05	-0.03	-0.15	-0.09
		DKd	0.25	0.28	0.26	0.39	0.05	0.04	0.15	0.15
		SW	-0.68	44.03	-1.02	5.64	0.98	1.32	3.28	4.41
	50	DK	0.27	0.14	0.26	0.18	0.03	0.04	0.11	0.13
		DKi	0.34	0.34	0.33	0.44	-0.04	-0.03	-0.15	-0.11
		DKd	0.28	0.25	0.26	0.32	0.02	0.04	0.08	0.12
		SW	-2.07	28.15	-0.51	3.79	2.37	0.81	7.89	2.69
	75	DK	0.28	0.11	0.28	0.15	0.02	0.02	0.05	0.05
		DKi	0.33	0.24	0.31	0.34	-0.03	-0.01	-0.09	-0.02
		DKd	0.28	0.16	0.29	0.21	0.02	0.01	0.05	0.04
		SW	-1.21	63.40	-0.02	2.20	1.51	0.32	5.03	1.06
60	25	DK	0.49	0.13	0.48	0.18	0.11	0.12	0.19	0.20
		DKi	0.65	0.31	0.64	0.41	-0.05	-0.04	-0.09	-0.07
		DKd	0.52	0.22	0.51	0.31	0.08	0.09	0.13	0.15
		SW	-5.65	144.55	-0.80	6.08	6.25	1.40	10.42	2.34
	50	DK	0.54	0.13	0.54	0.17	0.06	0.06	0.10	0.10
		DKi	0.63	0.27	0.61	0.35	-0.03	-0.01	-0.05	-0.02
		DKd	0.55	0.19	0.54	0.26	0.05	0.06	0.09	0.11
		SW	-0.85	35.60	-0.23	4.46	1.45	0.83	2.42	1.39
	75	DK	0.58	0.10	0.58	0.13	0.02	0.02	0.04	0.03
		DKi	0.63	0.21	0.63	0.27	-0.03	-0.03	-0.05	-0.04
		DKd	0.58	0.14	0.58	0.18	0.02	0.02	0.03	0.04
		SW	0.15	27.03	0.17	2.23	0.45	0.43	0.76	0.71
90	25	DK	0.73	0.11	0.73	0.15	0.17	0.17	0.18	0.19
		DKi	0.98	0.24	0.99	0.32	-0.08	-0.09	-0.09	-0.10
		DKd	0.76	0.18	0.76	0.25	0.14	0.14	0.15	0.16
		SW	5.98	236.19	-0.73	7.35	-5.08	1.63	-5.64	1.81
	50	DK	0.82	0.12	0.82	0.15	0.08	0.08	0.09	0.09
		DKi	0.95	0.23	0.94	0.28	-0.05	-0.04	-0.05	-0.04
		DKd	0.83	0.17	0.83	0.23	0.07	0.07	0.07	0.08
		SW	0.63	69.69	-0.30	4.36	0.27	1.20	0.30	1.33
	75	DK	0.87	0.09	0.86	0.12	0.03	0.04	0.04	0.04
		DKi	0.93	0.18	0.93	0.26	-0.03	-0.03	-0.04	-0.03
		DKd	0.87	0.13	0.87	0.17	0.03	0.03	0.03	0.03
		SW	1.64	25.71	0.29	2.48	-0.74	0.61	-0.83	0.68

Note: Gruber Saez estimator as given in equation (11) with Weber instrumentation ($k = 3$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

TABLE 5. Selected statistics of the simulated distribution of Gruber Saez estimator with base-year instrumentation ($k=0$), MA(1)

ETI	FE	TE	eti results				bias		relative bias	
			mean	sd	p50	iqr	mean	p50	mean	p50
30	25	DK	0.10	0.04	0.10	0.05	0.20	0.20	0.67	0.67
		DKi	0.08	0.07	0.08	0.10	0.22	0.22	0.73	0.73
		DKd	0.07	0.03	0.07	0.04	0.23	0.23	0.77	0.77
		SW	0.04	0.05	0.04	0.07	0.26	0.26	0.85	0.85
	50	DK	0.11	0.03	0.11	0.05	0.19	0.19	0.64	0.64
		DKi	0.08	0.07	0.08	0.10	0.22	0.22	0.74	0.74
		DKd	0.08	0.03	0.08	0.04	0.22	0.22	0.74	0.74
		SW	0.05	0.05	0.05	0.07	0.25	0.25	0.84	0.85
	75	DK	0.12	0.03	0.12	0.04	0.18	0.18	0.61	0.61
		DKi	0.08	0.05	0.08	0.07	0.22	0.22	0.73	0.74
		DKd	0.09	0.03	0.09	0.03	0.21	0.21	0.70	0.70
		SW	0.06	0.04	0.06	0.06	0.24	0.24	0.81	0.81
60	25	DK	0.17	0.03	0.17	0.05	0.43	0.43	0.72	0.72
		DKi	0.14	0.05	0.14	0.08	0.46	0.46	0.77	0.77
		DKd	0.11	0.03	0.11	0.04	0.49	0.49	0.81	0.81
		SW	0.07	0.05	0.07	0.06	0.53	0.53	0.88	0.88
	50	DK	0.18	0.03	0.18	0.04	0.42	0.42	0.69	0.69
		DKi	0.14	0.05	0.13	0.07	0.46	0.47	0.77	0.78
		DKd	0.13	0.03	0.13	0.03	0.47	0.47	0.79	0.79
		SW	0.08	0.04	0.08	0.06	0.52	0.52	0.86	0.87
	75	DK	0.20	0.03	0.20	0.04	0.40	0.40	0.67	0.66
		DKi	0.14	0.04	0.14	0.06	0.46	0.46	0.77	0.77
		DKd	0.14	0.02	0.14	0.03	0.46	0.46	0.77	0.77
		SW	0.09	0.04	0.09	0.05	0.51	0.51	0.85	0.85
90	25	DK	0.25	0.03	0.25	0.05	0.65	0.65	0.72	0.72
		DKi	0.20	0.05	0.20	0.07	0.70	0.70	0.78	0.78
		DKd	0.15	0.03	0.15	0.04	0.75	0.75	0.83	0.83
		SW	0.10	0.04	0.10	0.06	0.80	0.80	0.89	0.89
	50	DK	0.27	0.03	0.27	0.04	0.63	0.63	0.70	0.70
		DKi	0.21	0.05	0.21	0.06	0.69	0.69	0.77	0.77
		DKd	0.17	0.03	0.17	0.04	0.73	0.73	0.81	0.81
		SW	0.12	0.04	0.12	0.05	0.78	0.78	0.87	0.87
	75	DK	0.29	0.03	0.29	0.04	0.61	0.61	0.68	0.68
		DKi	0.22	0.04	0.22	0.05	0.68	0.68	0.76	0.76
		DKd	0.18	0.02	0.18	0.03	0.72	0.72	0.80	0.80
		SW	0.13	0.03	0.13	0.04	0.77	0.77	0.86	0.86

Note: Gruber Saez estimator as given in equation (11) with base-year instrumentation ($k = 0$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

TABLE 6. Selected statistics of the simulated distribution of Gruber Saez estimator with Weber instrumentation ($k=1$), MA(1)

ETI	FE	TE	eti results				bias		relative bias	
			mean	sd	p50	iqr	mean	p50	mean	p50
30	25	DK	0.27	0.14	0.27	0.18	0.03	0.03	0.11	0.12
		DKi	0.32	0.43	0.32	0.57	-0.02	-0.02	-0.06	-0.06
		DKd	0.11	0.24	0.10	0.31	0.19	0.20	0.64	0.68
		SW	0.45	4.39	0.12	2.53	-0.15	0.18	-0.50	0.62
	50	DK	0.27	0.12	0.26	0.16	0.03	0.04	0.11	0.12
		DKi	0.34	0.42	0.31	0.56	-0.04	-0.01	-0.13	-0.02
		DKd	0.16	0.19	0.16	0.25	0.14	0.14	0.47	0.47
		SW	0.50	12.07	0.18	1.68	-0.20	0.12	-0.66	0.40
	75	DK	0.28	0.09	0.28	0.12	0.02	0.02	0.07	0.08
		DKi	0.35	0.32	0.34	0.42	-0.05	-0.04	-0.18	-0.13
		DKd	0.22	0.13	0.21	0.17	0.08	0.09	0.27	0.29
		SW	0.43	1.35	0.23	1.16	-0.13	0.07	-0.42	0.24
60	25	DK	0.57	0.13	0.57	0.17	0.03	0.03	0.06	0.05
		DKi	0.51	0.34	0.50	0.44	0.09	0.10	0.15	0.16
		DKd	0.23	0.23	0.23	0.32	0.37	0.37	0.62	0.62
		SW	0.22	2.29	0.13	1.91	0.38	0.47	0.64	0.78
	50	DK	0.56	0.12	0.56	0.15	0.04	0.04	0.07	0.07
		DKi	0.56	0.31	0.55	0.40	0.04	0.05	0.07	0.08
		DKd	0.33	0.20	0.32	0.26	0.27	0.28	0.46	0.47
		SW	0.37	1.32	0.17	1.42	0.23	0.43	0.38	0.72
	75	DK	0.57	0.09	0.56	0.12	0.03	0.04	0.06	0.06
		DKi	0.63	0.25	0.62	0.34	-0.03	-0.02	-0.04	-0.03
		DKd	0.45	0.15	0.45	0.20	0.15	0.15	0.25	0.25
		SW	0.56	1.02	0.42	1.00	0.04	0.18	0.06	0.30
90	25	DK	0.87	0.12	0.87	0.16	0.03	0.03	0.03	0.03
		DKi	0.68	0.27	0.67	0.36	0.22	0.23	0.24	0.26
		DKd	0.34	0.23	0.34	0.31	0.56	0.56	0.62	0.62
		SW	0.19	1.43	0.12	1.78	0.71	0.78	0.78	0.87
	50	DK	0.87	0.11	0.87	0.15	0.03	0.03	0.04	0.03
		DKi	0.77	0.26	0.76	0.35	0.13	0.14	0.14	0.15
		DKd	0.51	0.20	0.51	0.26	0.39	0.39	0.43	0.43
		SW	0.47	1.14	0.39	1.36	0.43	0.51	0.48	0.57
	75	DK	0.87	0.09	0.87	0.11	0.03	0.03	0.03	0.03
		DKi	0.86	0.22	0.86	0.28	0.04	0.04	0.04	0.04
		DKd	0.69	0.15	0.69	0.20	0.21	0.21	0.23	0.23
		SW	0.62	0.75	0.54	0.84	0.28	0.36	0.31	0.40

Note: Gruber Saez estimator as given in equation (11) with Weber instrumentation ($k = 1$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

TABLE 7. Selected statistics of the simulated distribution of Gruber Saez estimator with Weber instrumentation ($k=2$), MA(1)

ETI	FE	TE	eti results				bias		relative bias	
			mean	sd	p50	iqr	mean	p50	mean	p50
30	25	DK	0.30	0.21	0.30	0.27	0.00	0.00	0.00	0.02
		DKi	0.34	0.46	0.33	0.65	-0.04	-0.03	-0.15	-0.11
		DKd	0.29	0.49	0.26	0.64	0.01	0.04	0.04	0.15
		SW	-1.66	49.35	-0.12	5.09	1.96	0.42	6.54	1.41
	50	DK	0.30	0.19	0.29	0.26	0.00	0.01	-0.01	0.03
		DKi	0.33	0.42	0.31	0.56	-0.03	-0.01	-0.11	-0.05
		DKd	0.30	0.40	0.29	0.54	0.00	0.01	0.00	0.04
		SW	11.54	273.14	-0.01	4.24	-11.24	0.31	-37.48	1.05
	75	DK	0.31	0.15	0.30	0.20	-0.01	0.00	-0.03	-0.01
		DKi	0.34	0.28	0.31	0.40	-0.04	-0.01	-0.12	-0.03
		DKd	0.31	0.26	0.30	0.36	-0.01	0.00	-0.05	0.00
		SW	-0.11	30.31	0.29	2.62	0.41	0.01	1.36	0.05
60	25	DK	0.61	0.19	0.60	0.26	-0.01	0.00	-0.01	0.00
		DKi	0.63	0.39	0.62	0.51	-0.03	-0.02	-0.04	-0.03
		DKd	0.61	0.37	0.61	0.49	-0.01	-0.01	-0.01	-0.02
		SW	0.71	34.70	0.27	5.07	-0.11	0.33	-0.19	0.55
	50	DK	0.61	0.17	0.60	0.21	-0.01	0.00	-0.01	-0.01
		DKi	0.62	0.34	0.60	0.45	-0.02	0.00	-0.03	-0.01
		DKd	0.61	0.33	0.59	0.44	-0.01	0.01	-0.02	0.01
		SW	12.48	340.61	-0.12	3.82	-11.88	0.72	-19.81	1.20
	75	DK	0.61	0.13	0.60	0.17	-0.01	0.00	-0.01	0.00
		DKi	0.62	0.25	0.61	0.34	-0.02	-0.01	-0.03	-0.01
		DKd	0.61	0.23	0.61	0.30	-0.01	-0.01	-0.02	-0.01
		SW	1.97	41.28	0.35	2.63	-1.37	0.25	-2.28	0.42
90	25	DK	0.91	0.17	0.91	0.23	-0.01	-0.01	-0.01	-0.01
		DKi	0.94	0.32	0.93	0.43	-0.04	-0.03	-0.04	-0.03
		DKd	0.92	0.32	0.92	0.45	-0.02	-0.02	-0.02	-0.03
		SW	-7.95	237.33	0.64	4.71	8.85	0.26	9.83	0.29
	50	DK	0.91	0.16	0.90	0.21	-0.01	0.00	-0.01	0.00
		DKi	0.93	0.28	0.92	0.36	-0.03	-0.02	-0.03	-0.02
		DKd	0.92	0.27	0.90	0.37	-0.02	0.00	-0.02	0.00
		SW	1.58	38.86	0.73	4.35	-0.68	0.17	-0.76	0.18
	75	DK	0.90	0.12	0.90	0.16	0.00	0.00	0.00	0.00
		DKi	0.94	0.23	0.92	0.30	-0.04	-0.02	-0.04	-0.02
		DKd	0.91	0.20	0.90	0.26	-0.01	0.00	-0.01	0.00
		SW	3.02	64.01	0.68	2.57	-2.12	0.22	-2.35	0.24

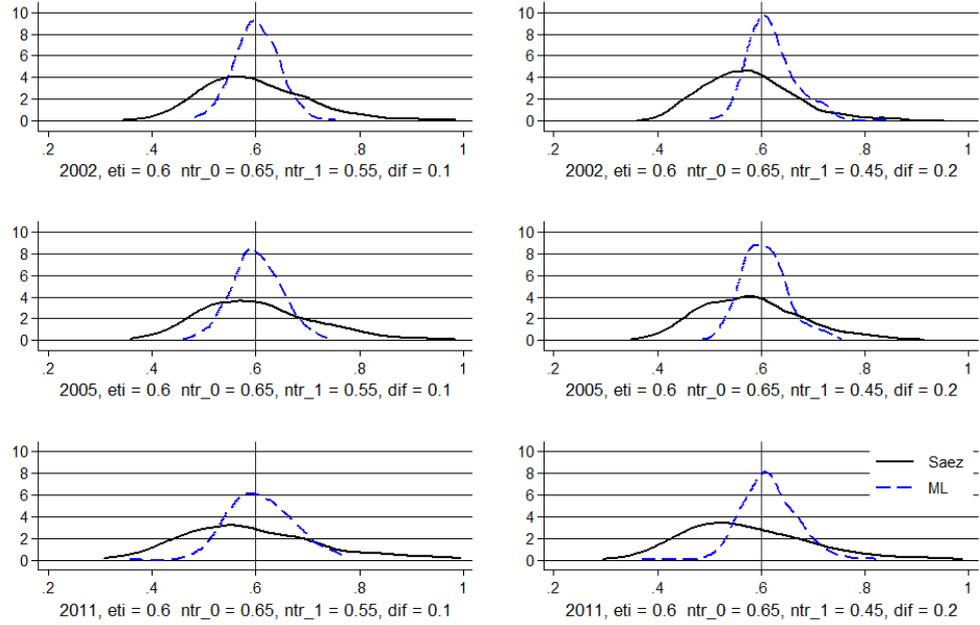
Note: Gruber Saez estimator as given in equation (11) with Weber instrumentation ($k = 2$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

TABLE 8. Selected statistics of the simulated distribution of Gruber Saez estimator with Weber instrumentation ($k=3$), MA(1)

ETI	FE	TE	eti results				bias		relative bias	
			mean	sd	p50	iqr	mean	p50	mean	p50
30	25	DK	0.30	0.22	0.30	0.29	0.00	0.00	-0.01	-0.01
		DKi	0.34	0.49	0.34	0.61	-0.04	-0.04	-0.13	-0.12
		DKd	0.32	0.46	0.29	0.57	-0.02	0.01	-0.05	0.04
		SW	-10.98	248.47	-2.07	8.67	11.28	2.37	37.60	7.91
	50	DK	0.31	0.20	0.30	0.25	-0.01	0.00	-0.02	-0.01
		DKi	0.32	0.44	0.30	0.57	-0.02	0.00	-0.07	0.00
		DKd	0.32	0.38	0.30	0.52	-0.02	0.00	-0.06	0.02
		SW	0.53	131.86	-1.71	6.36	-0.23	2.01	-0.76	6.71
	75	DK	0.31	0.14	0.31	0.19	-0.01	-0.01	-0.03	-0.03
		DKi	0.32	0.30	0.30	0.39	-0.02	0.00	-0.08	0.00
		DKd	0.32	0.25	0.30	0.34	-0.02	0.00	-0.05	-0.01
		SW	-1.48	23.82	-0.73	4.10	1.78	1.03	5.94	3.44
60	25	DK	0.60	0.19	0.61	0.25	0.00	-0.01	-0.01	-0.01
		DKi	0.62	0.41	0.60	0.52	-0.02	0.00	-0.03	0.00
		DKd	0.61	0.34	0.61	0.43	-0.01	-0.01	-0.02	-0.02
		SW	-9.67	266.28	-2.02	10.80	10.27	2.62	17.12	4.37
	50	DK	0.61	0.17	0.60	0.24	-0.01	0.00	-0.01	0.00
		DKi	0.64	0.36	0.60	0.47	-0.04	0.00	-0.06	0.00
		DKd	0.61	0.29	0.61	0.37	-0.01	-0.01	-0.02	-0.02
		SW	-33.76	701.93	-1.50	8.43	34.36	2.10	57.27	3.50
	75	DK	0.60	0.13	0.60	0.17	0.00	0.00	0.00	0.00
		DKi	0.62	0.26	0.61	0.34	-0.02	-0.01	-0.03	-0.02
		DKd	0.61	0.20	0.60	0.26	-0.01	0.00	-0.01	-0.01
		SW	-1.81	31.79	-0.75	4.27	2.41	1.35	4.02	2.25
90	25	DK	0.91	0.17	0.91	0.23	-0.01	-0.01	-0.01	-0.01
		DKi	0.92	0.33	0.91	0.43	-0.02	-0.01	-0.02	-0.01
		DKd	0.91	0.26	0.91	0.36	-0.01	-0.01	-0.01	-0.01
		SW	3.77	456.28	-2.53	11.68	-2.87	3.43	-3.19	3.81
	50	DK	0.91	0.16	0.91	0.23	-0.01	-0.01	-0.01	-0.01
		DKi	0.93	0.29	0.93	0.40	-0.03	-0.03	-0.04	-0.03
		DKd	0.92	0.23	0.92	0.30	-0.02	-0.02	-0.02	-0.02
		SW	0.20	169.16	-1.50	9.08	0.70	2.40	0.77	2.67
	75	DK	0.90	0.12	0.90	0.15	0.00	0.00	0.00	0.00
		DKi	0.93	0.23	0.93	0.30	-0.03	-0.03	-0.03	-0.03
		DKd	0.91	0.17	0.91	0.24	-0.01	-0.01	-0.02	-0.01
		SW	1.24	89.41	-0.54	4.57	-0.34	1.44	-0.38	1.60

Note: Gruber Saez estimator as given in equation (11) with Weber instrumentation ($k = 3$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

FIGURE 6. Effect of tax environment bunching estimator
 small kink (10 percentage points) large kink (20 percentage points)



Note: Saez and lognormal bunching estimators as described in appendix ???. 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: small kink: $ntr_0 = 0.65$, $ntr_1 = 0.55$, $dif = 0.1$; large kink: $ntr_0 = 0.65$, $ntr_1 = 0.45$, $dif = 0.2$. See table (17) and part (3) for details. Estimators: ETI_{Saez0} : original Saez estimator; ETI_{norm} : our preferred ML based ETI bunching estimator (equation 19 in appendix ??). See Table (10) for precise figures

TABLE 9. Selected statistics of the simulated distribution of the Bunching Estimator (ETI = 0.3)

		mean	sd	p50	iqr	relative bias		observation numbers			
						mean	median	bk	k	ak	total
large kink											
2002	ML	0.301	0.020	0.300	0.026	-0.004	-0.002	588	414	400	1401
2002	Saez	0.301	0.045	0.297	0.058	-0.004	0.010	39	444	29	512
2003	ML	0.300	0.019	0.300	0.027	0.001	0.001	548	394	395	1336
2003	Saez	0.302	0.049	0.298	0.066	-0.007	0.008	36	423	29	488
2004	ML	0.300	0.020	0.300	0.028	-0.001	-0.001	531	379	376	1286
2004	Saez	0.300	0.049	0.295	0.060	-0.001	0.015	35	408	28	471
2005	ML	0.300	0.020	0.300	0.026	-0.001	0.001	515	365	360	1240
2005	Saez	0.303	0.051	0.298	0.069	-0.009	0.008	34	392	26	452
2006	ML	0.300	0.021	0.300	0.028	-0.002	-0.001	497	351	345	1192
2006	Saez	0.304	0.050	0.301	0.068	-0.013	-0.002	33	377	25	434
2007	ML	0.300	0.021	0.300	0.029	-0.001	0.001	474	330	321	1125
2007	Saez	0.304	0.052	0.298	0.070	-0.014	0.005	31	355	23	409
2008	ML	0.301	0.022	0.300	0.030	-0.002	0.000	449	310	297	1056
2008	Saez	0.300	0.054	0.294	0.071	-0.001	0.021	30	333	22	384
2009	ML	0.302	0.025	0.302	0.034	-0.007	-0.006	374	250	234	858
2009	Saez	0.303	0.064	0.296	0.082	-0.011	0.012	24	269	18	310
2010	ML	0.304	0.025	0.303	0.034	-0.013	-0.010	364	244	226	835
2010	Saez	0.303	0.062	0.296	0.079	-0.012	0.012	23	262	17	302
2011	ML	0.303	0.027	0.302	0.034	-0.008	-0.007	347	229	212	788
2011	Saez	0.304	0.067	0.294	0.083	-0.013	0.019	22	246	16	284
2012	ML	0.310	0.030	0.308	0.041	-0.032	-0.026	317	208	189	714
2012	Saez	0.302	0.071	0.291	0.082	-0.005	0.031	20	223	14	258
small kink											
2002	ML	0.301	0.027	0.301	0.037	-0.005	-0.003	588	217	449	1254
2002	Saez	0.304	0.058	0.300	0.074	-0.015	0.000	39	252	33	323
2003	ML	0.300	0.028	0.299	0.038	0.000	0.005	548	206	437	1191
2003	Saez	0.305	0.061	0.300	0.083	-0.016	0.000	36	238	32	306
2004	ML	0.301	0.028	0.302	0.037	-0.003	-0.006	531	198	418	1148
2004	Saez	0.305	0.061	0.299	0.074	-0.018	0.004	35	230	30	296
2005	ML	0.300	0.028	0.301	0.036	-0.001	-0.002	515	191	401	1107
2005	Saez	0.305	0.063	0.298	0.085	-0.017	0.005	34	221	29	284
2006	ML	0.301	0.029	0.301	0.040	-0.004	-0.004	497	184	385	1065
2006	Saez	0.309	0.063	0.304	0.085	-0.031	-0.014	33	213	28	273
2007	ML	0.301	0.030	0.301	0.043	-0.003	-0.003	474	174	361	1008
2007	Saez	0.308	0.067	0.301	0.086	-0.028	-0.004	31	201	26	258
2008	ML	0.301	0.031	0.299	0.043	-0.004	0.002	449	163	336	948
2008	Saez	0.304	0.069	0.297	0.087	-0.012	0.011	30	188	25	243
2009	ML	0.306	0.036	0.305	0.050	-0.019	-0.017	374	133	267	774
2009	Saez	0.310	0.078	0.301	0.096	-0.032	-0.003	24	154	20	198
2010	ML	0.305	0.035	0.305	0.049	-0.018	-0.018	364	130	260	754
2010	Saez	0.306	0.077	0.298	0.099	-0.021	0.006	23	150	20	193
2011	ML	0.307	0.036	0.307	0.049	-0.022	-0.023	347	122	243	712
2011	Saez	0.307	0.075	0.299	0.096	-0.023	0.004	22	141	18	181
2012	ML	0.313	0.039	0.311	0.053	-0.042	-0.035	317	111	219	647
2012	Saez	0.307	0.084	0.302	0.106	-0.025	-0.005	20	128	16	165

Note: Saez and lognormal ML bunching estimators as described in part 4 and Aronsson et al., (2018). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: small kink: $ntr_0 = 0.65$, $ntr_1 = 0.55$, $dif = 0.1$; large kink: $ntr_0 = 0.65$, $ntr_1 = 0.45$, $dif = 0.2$. See table (17) and part 3 for details. Estimators: *Saez*: original Saez estimator as in Equation (17); *ML*: ML based ETI bunching estimator assuming lognormal income distribution as in Equation (19).

TABLE 10. Selected statistics of the simulated distribution of the Bunching Estimator (ETI = 0.6)

		mean	sd	p50	iqr	relative bias		observation numbers			
						mean	median	bk	k	ak	total
large kink											
2002	ML	0.619	0.048	0.616	0.059	-0.032	-0.026	482	557	231	1271
2002	Saez	0.583	0.096	0.570	0.122	0.029	0.049	31	576	18	624
2003	ML	0.618	0.046	0.612	0.060	-0.030	-0.019	459	556	242	1256
2003	Saez	0.589	0.098	0.579	0.126	0.019	0.035	30	574	18	622
2004	ML	0.615	0.045	0.611	0.059	-0.025	-0.018	442	528	229	1200
2004	Saez	0.581	0.096	0.570	0.126	0.032	0.051	29	547	17	593
2005	ML	0.616	0.047	0.611	0.059	-0.027	-0.019	425	505	217	1147
2005	Saez	0.588	0.107	0.573	0.134	0.019	0.044	28	522	16	566
2006	ML	0.617	0.045	0.614	0.057	-0.028	-0.023	409	482	204	1095
2006	Saez	0.585	0.103	0.577	0.133	0.025	0.039	27	498	15	540
2007	ML	0.616	0.049	0.611	0.061	-0.026	-0.019	385	448	188	1021
2007	Saez	0.583	0.110	0.572	0.140	0.028	0.047	25	463	14	502
2008	ML	0.615	0.049	0.612	0.063	-0.025	-0.020	361	413	171	946
2008	Saez	0.579	0.110	0.567	0.141	0.034	0.054	23	427	13	463
2009	ML	0.643	0.065	0.645	0.078	-0.071	-0.075	293	322	127	741
2009	Saez	0.576	0.121	0.561	0.154	0.040	0.065	19	332	10	360
2010	ML	0.641	0.066	0.643	0.088	-0.069	-0.072	285	311	123	719
2010	Saez	0.587	0.130	0.568	0.157	0.022	0.053	18	321	9	349
2011	ML	0.643	0.070	0.644	0.088	-0.072	-0.073	268	291	113	672
2011	Saez	0.584	0.132	0.563	0.167	0.027	0.061	17	300	9	326
2012	ML	0.633	0.072	0.639	0.095	-0.054	-0.066	243	258	99	600
2012	Saez	0.584	0.138	0.561	0.161	0.027	0.065	15	266	8	289
small kink											
2002	ML	0.601	0.046	0.601	0.064	-0.001	-0.002	482	300	310	1092
2002	Saez	0.604	0.109	0.592	0.137	-0.007	0.013	31	324	23	378
2003	ML	0.605	0.046	0.604	0.062	-0.008	-0.007	459	296	314	1069
2003	Saez	0.613	0.116	0.602	0.147	-0.021	-0.003	30	320	23	373
2004	ML	0.602	0.048	0.599	0.065	-0.003	0.001	442	282	299	1022
2004	Saez	0.604	0.117	0.588	0.137	-0.007	0.020	29	304	22	355
2005	ML	0.602	0.047	0.602	0.064	-0.004	-0.003	425	270	284	979
2005	Saez	0.607	0.116	0.599	0.158	-0.011	0.002	28	292	21	340
2006	ML	0.603	0.047	0.603	0.062	-0.006	-0.005	409	258	270	936
2006	Saez	0.603	0.115	0.594	0.155	-0.004	0.011	27	279	20	326
2007	ML	0.604	0.052	0.603	0.070	-0.007	-0.005	385	241	250	876
2007	Saez	0.610	0.124	0.595	0.167	-0.017	0.008	25	260	19	304
2008	ML	0.607	0.054	0.605	0.070	-0.011	-0.008	361	223	229	814
2008	Saez	0.607	0.127	0.595	0.173	-0.012	0.008	23	241	17	282
2009	ML	0.626	0.065	0.623	0.088	-0.044	-0.038	293	175	176	643
2009	Saez	0.602	0.141	0.588	0.188	-0.004	0.019	19	190	13	222
2010	ML	0.628	0.064	0.626	0.089	-0.047	-0.043	285	170	169	625
2010	Saez	0.616	0.148	0.595	0.185	-0.026	0.009	18	184	13	214
2011	ML	0.634	0.070	0.629	0.095	-0.056	-0.049	268	160	157	585
2011	Saez	0.610	0.149	0.593	0.193	-0.017	0.011	17	172	12	201
2012	ML	0.637	0.073	0.633	0.097	-0.062	-0.054	243	143	139	524
2012	Saez	0.614	0.157	0.590	0.206	-0.023	0.017	15	154	11	180

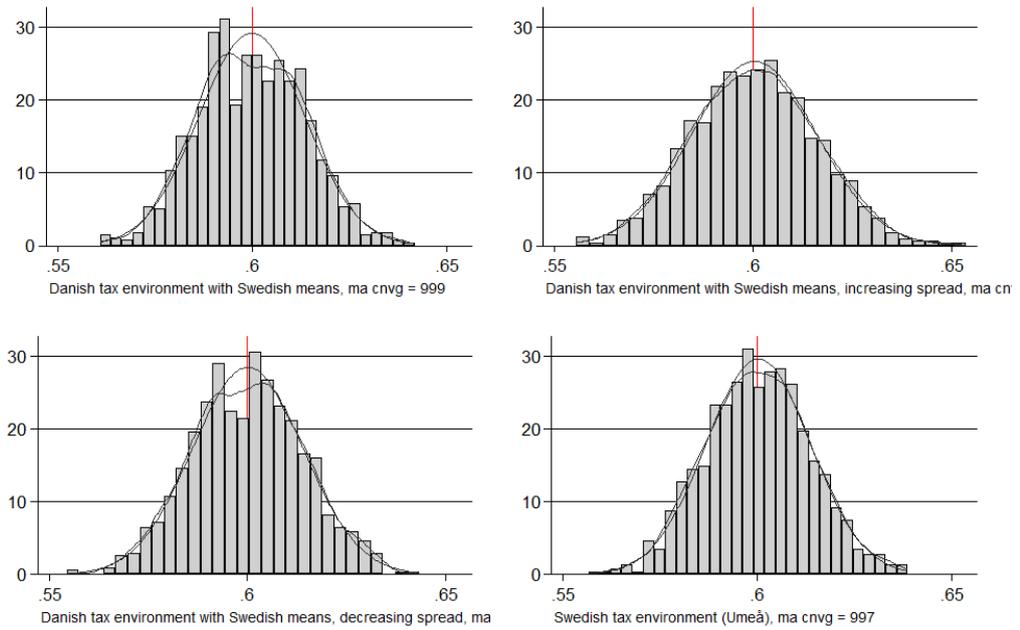
Note: Saez and lognormal ML bunching estimators as described in part 4 and Aronsson et al., (2018). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: small kink: $ntr_0 = 0.65$, $ntr_1 = 0.55$, $dif = 0.1$; large kink: $ntr_0 = 0.65$, $ntr_1 = 0.45$, $dif = 0.2$. See table (17) and part 3 for details. Estimators: *Saez*: original Saez estimator as in Equation (17); *ML*: ML based ETI bunching estimator assuming lognormal income distribution as in Equation (19).

TABLE 11. Selected statistics of the simulated distribution of the Bunching Estimator (ETI = 0.9)

		mean	sd	p50	iqr	relative bias		observation numbers			
						mean	median	bk	k	ak	total
large kink											
2002	ML	0.960	0.060	0.955	0.082	-0.066	-0.061	378	542	116	1036
2002	Saez	0.809	0.150	0.790	0.186	0.101	0.122	24	552	9	585
2003	ML	0.920	0.054	0.915	0.071	-0.022	-0.017	373	565	134	1072
2003	Saez	0.827	0.153	0.806	0.185	0.081	0.104	24	576	10	610
2004	ML	0.936	0.067	0.930	0.089	-0.040	-0.033	355	534	124	1014
2004	Saez	0.832	0.157	0.814	0.203	0.076	0.096	23	544	10	576
2005	ML	0.957	0.071	0.949	0.095	-0.063	-0.055	340	506	116	963
2005	Saez	0.816	0.151	0.799	0.202	0.093	0.113	22	515	9	546
2006	ML	0.977	0.072	0.972	0.094	-0.085	-0.080	326	479	108	913
2006	Saez	0.813	0.163	0.791	0.210	0.097	0.121	21	488	8	517
2007	ML	0.984	0.074	0.982	0.091	-0.094	-0.091	304	441	98	843
2007	Saez	0.824	0.163	0.802	0.200	0.084	0.109	19	449	7	476
2008	ML	0.981	0.081	0.975	0.107	-0.090	-0.084	282	401	88	770
2008	Saez	0.819	0.174	0.789	0.205	0.090	0.123	18	408	7	433
2009	ML	1.001	0.113	0.998	0.153	-0.113	-0.109	222	300	62	584
2009	Saez	0.814	0.207	0.778	0.249	0.096	0.136	14	305	5	324
2010	ML	1.007	0.118	1.000	0.156	-0.119	-0.111	214	290	59	564
2010	Saez	0.823	0.217	0.779	0.258	0.086	0.134	13	295	5	313
2011	ML	1.001	0.123	0.994	0.155	-0.112	-0.104	201	268	54	522
2011	Saez	0.809	0.221	0.770	0.253	0.101	0.144	12	272	4	289
2012	ML	0.989	0.142	0.983	0.184	-0.099	-0.092	180	234	46	460
2012	Saez	0.808	0.230	0.767	0.268	0.103	0.148	11	238	4	253
small kink											
2002	ML	0.924	0.072	0.921	0.094	-0.027	-0.023	378	314	198	889
2002	Saez	0.894	0.172	0.872	0.219	0.006	0.031	24	329	15	369
2003	ML	0.908	0.068	0.905	0.097	-0.009	-0.006	373	320	212	904
2003	Saez	0.900	0.180	0.872	0.212	-0.001	0.031	24	336	16	376
2004	ML	0.921	0.077	0.915	0.096	-0.023	-0.017	355	304	199	858
2004	Saez	0.909	0.171	0.892	0.222	-0.010	0.009	23	319	15	357
2005	ML	0.933	0.081	0.932	0.111	-0.037	-0.035	340	288	188	816
2005	Saez	0.898	0.185	0.873	0.233	0.002	0.030	22	303	14	339
2006	ML	0.939	0.081	0.939	0.107	-0.044	-0.043	326	273	177	776
2006	Saez	0.886	0.193	0.863	0.238	0.015	0.041	21	287	14	322
2007	ML	0.950	0.089	0.944	0.128	-0.056	-0.049	304	253	162	719
2007	Saez	0.903	0.197	0.880	0.245	-0.004	0.022	19	266	12	298
2008	ML	0.955	0.094	0.949	0.128	-0.061	-0.054	282	232	147	660
2008	Saez	0.907	0.214	0.873	0.243	-0.008	0.030	18	243	11	272
2009	ML	0.973	0.111	0.968	0.148	-0.081	-0.075	222	176	107	505
2009	Saez	0.906	0.247	0.868	0.285	-0.006	0.036	14	185	8	207
2010	ML	0.974	0.108	0.967	0.147	-0.083	-0.075	214	171	103	488
2010	Saez	0.920	0.253	0.867	0.281	-0.022	0.036	13	179	8	200
2011	ML	0.967	0.115	0.956	0.149	-0.074	-0.062	201	158	95	454
2011	Saez	0.899	0.262	0.862	0.317	0.001	0.043	12	165	8	185
2012	ML	0.964	0.120	0.956	0.151	-0.071	-0.063	180	139	83	401
2012	Saez	0.917	0.281	0.872	0.334	-0.019	0.031	11	146	6	163

Note: Saez and lognormal ML bunching estimators as described in part 4 and Aronsson et al., (2018). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: small kink: $ntr_0 = 0.65$, $ntr_1 = 0.55$, $dif = 0.1$; large kink: $ntr_0 = 0.65$, $ntr_1 = 0.45$, $dif = 0.2$. See table (17) and part 3 for details. Estimators: *Saez*: original Saez estimator as in Equation (17); *ML*: ML based ETI bunching estimator assuming lognormal income distribution as in Equation (19).

FIGURE 7. Effect of tax environment indirect inference estimator MA



Note: Indirect inference estimator as described in part (4). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment given in graph note. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See Table (17) and Section (3) for details. See Table (13) for precise figures.

TABLE 12. Selected statistics of the simulated distribution of the Indirect Inference estimator, AR(1)

eti	FE%	tax environment	ETI results				relative bias		converged	reps
			mean	sd	p50	iqr	mean	p50		
	25	DK	0.300	0.008	0.300	0.010	0.001	0.001	988	1000
		DKi	0.300	0.010	0.300	0.013	0.000	-0.001	983	1000
		DKd	0.300	0.009	0.300	0.011	0.001	0.001	991	1000
		SW	0.300	0.008	0.300	0.011	0.001	0.000	987	1000
30	50	DK	0.300	0.008	0.300	0.011	0.000	0.000	991	1000
		DKi	0.300	0.010	0.300	0.014	0.000	-0.001	991	1000
		DKd	0.300	0.009	0.300	0.012	0.000	-0.001	990	1000
		SW	0.300	0.009	0.300	0.011	0.000	-0.001	988	1000
	75	DK	0.300	0.009	0.300	0.012	0.000	-0.001	992	1000
		DKi	0.300	0.010	0.300	0.015	0.001	-0.001	987	1000
		DKd	0.300	0.009	0.300	0.013	0.000	0.000	993	1000
		SW	0.300	0.009	0.300	0.013	0.000	-0.001	987	1000
	25	DK	0.600	0.014	0.600	0.019	0.000	0.000	998	1000
		DKi	0.600	0.016	0.600	0.020	0.000	0.000	997	1000
		DKd	0.600	0.014	0.599	0.019	0.000	-0.001	998	1000
		SW	0.600	0.014	0.600	0.018	0.000	-0.001	999	1000
60	50	DK	0.600	0.014	0.600	0.021	0.001	0.000	999	1000
		DKi	0.600	0.016	0.600	0.022	0.001	0.000	997	1000
		DKd	0.600	0.015	0.600	0.021	0.001	-0.001	995	1000
		SW	0.600	0.014	0.600	0.019	0.001	0.000	992	1000
	75	DK	0.600	0.017	0.599	0.022	0.000	-0.001	995	1000
		DKi	0.599	0.018	0.599	0.024	-0.001	-0.001	995	1000
		DKd	0.600	0.017	0.599	0.023	0.000	-0.002	993	1000
		SW	0.600	0.017	0.600	0.023	0.000	0.000	991	1000
	25	DK	0.900	0.022	0.899	0.029	0.000	-0.001	1000	1000
		DKi	0.901	0.025	0.899	0.033	0.001	-0.001	999	1000
		DKd	0.901	0.022	0.900	0.029	0.001	0.000	998	1000
		SW	0.901	0.021	0.900	0.028	0.001	0.000	997	1000
90	50	DK	0.900	0.024	0.899	0.031	0.000	-0.001	999	1000
		DKi	0.900	0.026	0.899	0.035	0.000	-0.001	995	1000
		DKd	0.900	0.024	0.900	0.031	0.000	0.000	995	1000
		SW	0.900	0.024	0.900	0.031	0.000	0.000	996	1000
	75	DK	0.899	0.031	0.899	0.042	-0.001	-0.001	996	1000
		DKi	0.900	0.031	0.900	0.041	0.000	0.000	992	1000
		DKd	0.899	0.030	0.898	0.041	-0.001	-0.002	994	1000
		SW	0.899	0.028	0.899	0.038	-0.001	-0.001	994	1000

Note: Indirect inference estimator as described in part (4). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

TABLE 13. Selected statistics of the simulated distribution of the Indirect Inference estimator, MA(1)

eti	FE%	tax environment	ETI results				relative bias		converged	reps
			mean	sd	p50	iqr	mean	p50		
	25	DK	0.300	0.008	0.300	0.011	-0.001	-0.001	988	1000
		DKi	0.300	0.009	0.299	0.013	-0.001	-0.002	985	1000
		DKd	0.300	0.008	0.300	0.011	-0.001	-0.001	988	1000
		SW	0.300	0.008	0.300	0.010	-0.001	-0.001	988	1000
30	50	DK	0.300	0.008	0.300	0.010	0.000	0.000	992	1000
		DKi	0.300	0.009	0.300	0.012	0.000	0.001	989	1000
		DKd	0.300	0.008	0.300	0.011	0.001	0.001	991	1000
		SW	0.300	0.008	0.300	0.011	0.000	0.001	991	1000
	75	DK	0.300	0.008	0.300	0.011	-0.001	-0.001	990	1000
		DKi	0.300	0.010	0.300	0.014	0.000	0.000	988	1000
		DKd	0.299	0.009	0.299	0.012	-0.002	-0.002	991	1000
		SW	0.300	0.009	0.300	0.012	-0.001	-0.001	987	1000
	25	DK	0.600	0.014	0.600	0.018	0.000	0.000	998	1000
		DKi	0.600	0.015	0.600	0.021	0.000	0.000	997	1000
		DKd	0.600	0.014	0.600	0.019	0.000	0.000	998	1000
		SW	0.600	0.013	0.600	0.018	0.000	-0.001	998	1000
60	50	DK	0.600	0.014	0.600	0.020	0.000	0.000	999	1000
		DKi	0.600	0.016	0.600	0.021	0.001	0.000	997	1000
		DKd	0.600	0.014	0.600	0.020	0.000	0.001	999	1000
		SW	0.600	0.013	0.600	0.018	0.000	0.000	997	1000
	75	DK	0.600	0.015	0.600	0.020	0.000	0.000	1000	1000
		DKi	0.600	0.017	0.601	0.023	0.001	0.001	999	1000
		DKd	0.600	0.016	0.600	0.022	0.000	0.000	1000	1000
		SW	0.600	0.015	0.601	0.020	0.001	0.001	998	1000
	25	DK	0.899	0.020	0.899	0.026	-0.001	-0.001	1000	1000
		DKi	0.901	0.023	0.900	0.030	0.001	0.000	998	1000
		DKd	0.899	0.021	0.899	0.028	-0.001	-0.001	1000	1000
		SW	0.900	0.020	0.899	0.026	0.000	-0.001	999	1000
90	50	DK	0.900	0.022	0.899	0.029	-0.001	-0.001	1000	1000
		DKi	0.900	0.024	0.900	0.033	0.000	0.000	1000	1000
		DKd	0.900	0.022	0.901	0.030	0.000	0.001	1000	1000
		SW	0.900	0.021	0.901	0.028	0.001	0.001	1000	1000
	75	DK	0.900	0.027	0.900	0.037	0.000	-0.001	1000	1000
		DKi	0.899	0.028	0.899	0.037	-0.001	-0.001	1000	1000
		DKd	0.900	0.026	0.900	0.036	0.000	0.000	1000	1000
		SW	0.900	0.025	0.899	0.035	0.000	-0.001	1000	1000

Note: Indirect inference estimator as described in part (4). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See table (17) and part 3 for details.

TABLE 14. Selected statistics of the simulated distribution of Gruber Saez estimator with Weber instrumentation ($k=3$), robustness tests

error	prop	dist	eti results				bias		relative bias	
			mean	sd	p50	iqr	mean	p50	mean	p50
AR	50	n	0.526	0.139	0.531	0.194	0.074	0.069	0.124	0.116
	75	n	0.531	0.132	0.533	0.185	0.069	0.067	0.116	0.111
	100	s(5)	0.542	0.127	0.542	0.170	0.058	0.058	0.097	0.097
	100	s(10)	0.537	0.125	0.532	0.163	0.063	0.068	0.106	0.113
MA	50	n	0.589	0.194	0.587	0.261	0.011	0.013	0.018	0.021
	75	n	0.598	0.187	0.595	0.257	0.002	0.005	0.003	0.008
	100	s(5)	0.607	0.169	0.609	0.236	-0.007	-0.009	-0.012	-0.015
	100	s(10)	0.593	0.168	0.587	0.228	0.007	0.013	0.012	0.022

Note: Gruber Saez estimator as given in equation 11, with Weber instrumentation ($k=3$). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. The ETI α is set to 0.6 for those tax units that respond, and to 0 for all other tax units. The share of the variance attributed to the permanent effect ϕ is set to 0.5. Prop: proportion of tax units that react to taxation. Dist: distribution of unobserved component; n \equiv log-normal; s(i) \equiv log-Student's t-distribution with i degrees of freedom. Tax environment: DK Danish tax variability with Swedish tax rate means. See table (17) and part (3) for details.

TABLE 15. Selected statistics of the simulated distribution of the Bunching Estimator (ETI = 0.6), robustness tests

	prop	dist	mean	sd	p50	iqr	relative bias		observation numbers				
							mean	median	bk	k	ak	total	
large kink													
2002	ML	50	n	0.276	0.021	0.275	0.028	0.540	0.541	474	298	306	1078
		75	n	0.432	0.028	0.431	0.037	0.279	0.282	478	432	272	1183
		100	s(5)	0.613	0.045	0.608	0.059	-0.021	-0.014	429	462	183	1074
		100	s(10)	0.616	0.044	0.612	0.057	-0.026	-0.020	462	523	215	1200
	Saez	50	n	0.275	0.049	0.269	0.061	0.542	0.552	31	322	23	375
		75	n	0.424	0.069	0.418	0.088	0.293	0.304	31	453	21	505
		100	s(5)	0.572	0.098	0.559	0.124	0.046	0.068	27	477	14	518
		100	s(10)	0.582	0.104	0.568	0.125	0.031	0.053	29	540	16	586
2011	ML	50	n	0.279	0.032	0.277	0.043	0.535	0.538	266	154	156	575
		75	n	0.449	0.044	0.448	0.056	0.251	0.254	267	222	135	624
		100	s(5)	0.615	0.071	0.620	0.096	-0.025	-0.033	219	231	91	542
		100	s(10)	0.636	0.071	0.641	0.096	-0.060	-0.068	248	267	104	619
	Saez	50	n	0.272	0.068	0.262	0.095	0.547	0.563	17	166	12	195
		75	n	0.423	0.098	0.410	0.130	0.295	0.317	17	233	10	260
		100	s(5)	0.585	0.153	0.556	0.179	0.025	0.073	14	239	7	259
		100	s(10)	0.587	0.134	0.572	0.169	0.022	0.047	15	275	8	298
small kink													
2002	ML	50	n	0.291	0.030	0.289	0.042	0.516	0.518	478	167	350	995
		75	n	0.444	0.038	0.443	0.050	0.260	0.262	481	236	333	1049
		100	s(5)	0.605	0.049	0.603	0.066	-0.008	-0.004	429	253	251	933
		100	s(10)	0.602	0.046	0.601	0.060	-0.004	-0.002	462	283	289	1034
	Saez	50	n	0.295	0.067	0.289	0.089	0.508	0.519	31	193	26	250
		75	n	0.447	0.087	0.440	0.109	0.254	0.266	31	261	25	317
		100	s(5)	0.606	0.120	0.590	0.148	-0.009	0.017	27	273	19	319
		100	s(10)	0.601	0.109	0.589	0.140	-0.002	0.019	29	306	22	358
2011	ML	50	n	0.298	0.042	0.297	0.055	0.504	0.505	267	88	179	534
		75	n	0.457	0.054	0.457	0.073	0.238	0.239	268	123	169	560
		100	s(5)	0.631	0.077	0.627	0.104	-0.052	-0.045	219	127	125	472
		100	s(10)	0.631	0.071	0.626	0.097	-0.051	-0.044	248	146	145	539
	Saez	50	n	0.296	0.087	0.286	0.108	0.507	0.523	17	102	14	132
		75	n	0.452	0.120	0.436	0.153	0.247	0.273	17	137	13	166
		100	s(5)	0.616	0.179	0.590	0.198	-0.027	0.017	14	137	10	160
		100	s(10)	0.621	0.157	0.596	0.198	-0.034	0.006	15	158	11	184

Saez and log-normal ML bunching estimators as described in part 4 and Aronsson et al., (2018). 1,000 replications with 10,000 individual observations over 12 years each. The ETI α is set to 0.6 for those tax units that respond, and to 0 for all other tax units. Prop: proportion of tax units that react to taxation. Dist: distribution of unobserved component; n \equiv log-normal; s(i) \equiv log-Student's t-distribution with i degrees of freedom. Tax environments: small kink: $ntr_0 = 0.65$, $ntr_1 = 0.55$, $dif = 0.1$; large kink: $ntr_0 = 0.65$, $ntr_1 = 0.45$, $dif = 0.2$. See table (17) and part (3) for details.

TABLE 16. Selected statistics of the simulated distribution of the Indirect Inference estimator, robustness tests

error	prop	dist	ETI results				relative bias		converged	reps
			mean	sd	p50	iqr	mean	p50		
AR	50	n	0.276	0.033	0.275	0.045	-0.541	-0.541	943	1000
	75	n	0.429	0.040	0.427	0.054	-0.285	-0.288	937	1000
	100	s(5)	0.571	0.049	0.571	0.063	-0.048	-0.048	540	1000
	100	s(10)	0.586	0.049	0.585	0.068	-0.023	-0.025	867	1000
MA	50	n	0.276	0.031	0.275	0.044	-0.539	-0.542	999	1000
	75	n	0.429	0.036	0.427	0.050	-0.286	-0.288	996	1000
	100	s(5)	0.578	0.049	0.578	0.067	-0.037	-0.037	887	1000
	100	s(10)	0.583	0.043	0.582	0.055	-0.029	-0.029	994	1000

Note: Indirect inference estimator as described in part (4). 1,000 replications with 10,000 individual observations over 12 years each, facing the tax environment in column 3. The ETI α is set to 0.6 for those tax units that respond, and to 0 for all other tax units. The share of the variance attributed to the permanent effect ϕ is set to 0.5. Prop: proportion of tax units that react to taxation. Dist: distribution of unobserved component; n \equiv log-normal; s(i) \equiv log-Student's t-distribution with i degrees of freedom. Tax environment: DK Danish tax variability with Swedish tax rate means. See table (17) and part (3) for details..

REFERENCES

- Aronsson, T., K. Jenderny, and G. Lanot. *Alternative Parametric Bunching Estimators of the ETI*. Umeå School of Business and Economics, Umeå Economic Studies 956, 2018.
- Auten, G. and R. Carroll. “The Effect Of Income Taxes On Household Income”. In: *The Review of Economics and Statistics* 81.4 (1999), pp. 681–693.
- Bastani, S. and H. Selin. *Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule*. CESifo Working Paper Series 3865. CESifo Group Munich, 2012. URL: https://ideas.repec.org/p/ces/ceswps/_3865.html.
- “Bunching and non-bunching at kink points of the Swedish tax schedule”. In: *Journal of Public Economics* 109.C (2014), pp. 36–49.
- Blundell, R. and T. MaCurdy. “Chapter 27 - Labor Supply: A Review of Alternative Approaches”. In: ed. by O. C. Ashenfelter and D. Card. Vol. 3, Part A. *Handbook of Labor Economics*. Elsevier, 1999, pp. 1559–1695. DOI: [http://dx.doi.org/10.1016/S1573-4463\(99\)03008-4](http://dx.doi.org/10.1016/S1573-4463(99)03008-4).
- Blundell, R., T. MaCurdy, and C. Meghir. “Chapter 69 Labor Supply Models: Unobserved Heterogeneity, Nonparticipation and Dynamics”. In: ed. by J. J. Heckman and E. E. Leamer. Vol. 6, Part A. *Handbook of Econometrics*. Elsevier, 2007, pp. 4667–4775. DOI: [http://dx.doi.org/10.1016/S1573-4412\(07\)06069-2](http://dx.doi.org/10.1016/S1573-4412(07)06069-2).
- Browning, M. and M. Ejrnæs. “Heterogeneity in the Dynamics of Labor Earnings”. In: *Annual Review of Economics* 5.1 (2013), pp. 219–245.
- Browning, M., M. Ejrnæs, and J. Alvarez. “Modelling Income Processes with Lots of Heterogeneity”. In: *Review of Economic Studies* 77.4 (2010), pp. 1353–1381.
- Carroll, R. *Do Taxpayers Really Respond to Changes in Tax Rates? Evidence from the 1993 Tax Act*. OTA Working Paper 79. U.S. Department of the Treasury, 1998.
- Chetty, R. “Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply”. In: *Econometrica* 80.3 (2012), pp. 969–1018.
- Chetty, R., J. N. Friedman, and E. Saez. “Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings”. In: *American Economic Review* 103.7 (2013), pp. 2683–2721. DOI: [10.1257/aer.103.7.2683](https://doi.org/10.1257/aer.103.7.2683).
- Chetty, R., J. Friedman, T. Olsen, and L. Pistaferri. “Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from

- Danish Tax Records". In: *Quarterly Journal of Economics* 126.2 (2011), pp. 749–804.
- Doerrenberg, P., A. Peichl, and S. Siegloch. "The Elasticity of Taxable Income in the Presence of Deduction Possibilities". In: *Journal of Public Economics* 151 (2017), pp. 41–55.
- Einav, L., A. Finkelstein, and P. Schrimpf. "The Response of Drug Expenditure to Nonlinear Contract Design: Evidence from Medicare Part D *". In: *The Quarterly Journal of Economics* 130.2 (2015), pp. 841–899. DOI: 10.1093/qje/qjv005.
- "Bunching at the kink: Implications for spending responses to health insurance contracts". In: *Journal of Public Economics* 146 (2017), pp. 27–40. ISSN: 0047-2727. DOI: <https://doi.org/10.1016/j.jpubeco.2016.11.011>. URL: <http://www.sciencedirect.com/science/article/pii/S0047272716301670>.
- Ericson, P., L. Flood, and N. Islam. "Taxes, wages and working hours". In: *Empirical Economics* 49.2 (2015), pp. 503–535. ISSN: 0377-7332. DOI: 10.1007/s00181-014-0880-0. URL: <http://dx.doi.org/10.1007/s00181-014-0880-0>.
- Feenberg, D. R. and J. M. Poterba. "Income Inequality and the Incomes of Very High-Income Taxpayers: Evidence from Tax Returns". In: *Tax Policy and the Economy, vol.7*. Ed. by J. M. Poterba. MIT Press, 1993, pp. 145–77.
- Feldstein, M. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act". In: *Journal of Political Economy* 103.3 (1995), pp. 551–72.
- Gouriéroux, C., A. Monfort, and E. Renault. "Indirect inference". In: *Journal of Applied Econometrics* 8.S1 (1993), S85–S118. ISSN: 1099-1255. DOI: 10.1002/jae.3950080507. URL: <http://dx.doi.org/10.1002/jae.3950080507>.
- Gruber, J. and E. Saez. "The elasticity of taxable income: evidence and implications". In: *Journal of Public Economics* 84.1 (2002), pp. 1–32.
- Hargaden, E. P. *Taxpayer Responses over the Cycle: Evidence from Irish Notches*. Discussion Papers. University of Michigan Department of Economics, 2015.
- Jiang, W. and B. Turnbull. "The indirect method: Inference based on intermediate statistics - a synthesis and examples". In: *Statistical Science* 19.2 (2004), pp. 239–263.
- Kleven, H. J. and M. Waseem. "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan". In: *The Quarterly Journal of Economics* (2013). DOI: 10.1093/qje/qjt004.

- Kleven, H. J. “Bunching”. In: *Annual Review of Economics* 8.1 (2016), pp. 435–464. DOI: 10.1146/annurev-economics-080315-015234.
- Kleven, H. J. and E. A. Schultz. “Estimating Taxable Income Responses Using Danish Tax Reforms”. In: *American Economic Journal: Economic Policy* 6.4 (2014), pp. 271–301.
- Kopczuk, W. *Tax Bases, Tax Rates and the Elasticity of Reported Income*. Working Paper 10044. National Bureau of Economic Research, 2003.
- “Tax bases, tax rates and the elasticity of reported income”. In: *Journal of Public Economics* 89.11-12 (2005), pp. 2093–2119.
- Low, H. and L. Pistaferri. “Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off”. In: *American Economic Review* 105.10 (2015), pp. 2986–3029.
- Neisser, C. *The elasticity of taxable income: A meta-regression analysis*. ZEW Discussion Papers 17-032. Center for European Economic Research, 2017.
- Nenov, P. T. “Regional reallocation and housing markets in a model of frictional migration”. In: *Review of Economic Dynamics* 18.4 (2015), pp. 863–880. ISSN: 1094-2025. DOI: <http://dx.doi.org/10.1016/j.red.2015.08.002>. URL: <http://www.sciencedirect.com/science/article/pii/S109420251500054X>.
- Saez, E. *The Effect of Marginal Tax Rates on Income: A Panel Study of 'Bracket Creep'*. NBER Working Papers 7367. 1999.
- “The effect of marginal tax rates on income: a panel study of 'bracket creep'”. In: *Journal of Public Economics* 87.5-6 (2003), pp. 1231–1258.
- “Do Taxpayers Bunch at Kink Points?” In: *American Economic Journal: Economic Policy* 2.3 (2010), pp. 180–212.
- Saez, E., J. Slemrod, and S. H. Giertz. “The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review”. In: *Journal of Economic Literature* 50.1 (2012), pp. 3–50.
- Slemrod, J. “Buenas notches: lines and notches in tax system design”. In: *eJournal of Tax Research* 11.3 (2013), pp. 259–283.
- Weber, C. E. “Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences”. In: *Journal of Public Economics* 117.C (2014), pp. 90–103.

APPENDIX A. DETAILS OF SIMPLIFIED SWEDISH TAX SYSTEM

TABLE 17. Rates and Threshold levels of the Simplified Tax Systems.

year	tax environment								kink earnings
	SW		DK		DKi		DKd		
	τ_0^c	τ_1^c	τ_0^c	τ_1^c	τ_0^c	τ_1^c	τ_0^c	τ_1^c	
2002	.677	.477	.652	.395	.627	.530	.727	.395	252000
2003	.677	.477	.652	.395	.627	.530	.727	.395	273800
2004	.677	.477	.652	.395	.627	.530	.714	.395	284300
2005	.669	.469	.627	.446	.652	.519	.689	.436	291800
2006	.669	.469	.627	.441	.652	.509	.689	.441	298600
2007	.669	.469	.627	.436	.652	.509	.689	.446	306000
2008	.669	.469	.664	.467	.664	.498	.664	.467	316700
2009	.669	.469	.664	.498	.664	.467	.664	.498	328800
2010	.669	.469	.689	.509	.689	.446	.652	.509	367600
2011	.669	.469	.689	.509	.689	.441	.652	.509	372100
2012	.669	.469	.689	.519	.689	.436	.652	.519	383000
2013	.669	.469	.714	.530	.714	.395	.627	.530	401100
2014	.669	.469	.727	.530	.727	.395	.627	.530	413200
2015	.669	.469	.727	.530	.727	.395	.627	.530	433200

Note: The table describes the net of tax rates before, τ_0^c , and after, τ_1^c , the earnings limit for τ_1^c over time (in nominal SEK; 400000 SEK \approx 45000 USD). Tax environments: DK Danish tax variability with Swedish tax rate means, DKi same tax rates, ordered with increasing spread, DKd decreasing spread, SW simplified tax system Umeå, Sweden. See part 3 for details.

APPENDIX B. SELECTED ESTIMATES IN THE LITERATURE

TABLE 18. Selected IV Estimates in the Literature

paper	obs		ETI	se	ase	t-stat	cutoff	k	δ	yy	income
	total	annual									
Weber	24576	6044	0.858	0.311	0.24	2.8	\$ 10k	2	2	6	ti
	25097	6157	0.475	0.268	0.21	1.8	\$ 10.5k	2	2	6	bi
GS	59199	6578 ^a	0.611	0.144	0.12	4.2	\$ 10k	0	3	9	ti
	69129	7681 ^a	0.170	0.106	0.09	1.6	\$ 10k	0	3	9	bi
DPS	2713842	387692 ^a	0.337	0.023	0.14	14.91	€ 10k	1	1	7	ti
	2713842	387692 ^a	0.144	0.014	0.09	10.44	€ 10k	1	1	7	bi
	2261535	376923 ^a	0.406	0.021	0.13	19.7	€ 10k	2	2	6	ti
	2261535	376923 ^a	0.139	0.014	0.08	10.17	€ 10k	2	2	6	bi
KS	29398652	1547297 ^a	0.064	0.003	0.04	21.3	DKK 0	0	3	19	ti
	29540762	1554777 ^a	0.052	0.003	0.04	17.3	DKK 0	0	3	19	bi

Note: The table describes the standard deviations of different regression-based ETI studies. Weber: Weber, 2014, GS: Gruber and Saez, 2002, DPS: Doerrenberg et al., 2017, KS: Kleven and Schultz, 2014. obs: total and annual observation numbers, ^a: annual average, calculated by us based on total observations, ETI: estimated ETI parameter, se: standard error of estimation (we assume that all standard errors were clustered at the individual level), ase: adjusted standard error (corresponding to the number of annual individual observations in our Monte Carlo simulation), t-stat: t-statistic, cutoff: income cutoff, k : instrument lag before base year, δ : year difference between base year and end year, yy : panel length (number of year differences), income: income concept (ti: taxable income. bi: broad income).

TABLE 19. Selected Bunching Estimates in the Literature

paper	annual obs	ETI	data	years	sub-group
Saez	80000 to 200000	0.162	IRS	1988-2002	all tax filers
		0.209		1960-1969	all tax filers
CFOP	3811602 ^a 230758 ^a	0.001	DKRD	1994-2001	wage earners
		0.240			self-employed
BS	6083914 ^{a,b} 796922 ^a	0.004	SERD	1999-2005	full population
		0.024		2000-2008	self-employed

Note: Saez: Saez, 2010, CFOP: Chetty et al., 2011, BS: Bastani and Selin, 2014; IRS: annual cross-sections, Individual Public Use Tax Files; DKRD: Danish register data; SERD: Swedish register data. ^a annual average, calculated by us based on total observations, ^b reported in working paper version Bastani and Selin, 2012.

Table 18 reports point estimates and standard errors for several IV regression analyses in the literature. Correcting for the number of unique individual observations, we can adjust the reported standard errors for the ETI parameter such that

they correspond to the sample size of our Monte Carlo simulations (10000 annual individual observations). These adjusted standard errors (ase) are reported in the sixth column of Table 18, and range between 0.04 and 0.24, depending on the dataset used and specification details.

We can compare these adjusted standard errors to the standard deviations of the IV estimators that we find in our simulations. In the Weber specifications ($k > 0$), our standard deviations of the estimators range roughly between 0.1 and 0.2 for the most informative tax environment, and increase in other tax environments. When we reproduce GSs approach, our standard deviations of the point estimate do typically not exceed 5. The estimates reported here show that the variability of the estimators produced by our simulations is of comparable size as the range of empirically observed results.

Table 19 reports estimates for several bunching analyses in the literature. It shows that bunching estimates for Denmark and the US are considerably smaller than the IV regression estimates reported in Table 18. These comparisons are of course illustrative, as the identifying populations are not identical.

APPENDIX C. AUXILIARY STATISTICS FOR INDIRECT INFERENCE

The first set of auxiliary statistics captures means of the earnings distribution and earnings growth. We characterize the mean log earnings below and above the kink M_{level} , $M_{\text{level} > \text{kink}}$ and the mean earnings growth M_{growth} . To calculate the first two quantities we smooth the indicator function $1_{[x < 0]}$ using a simple logit cumulative distribution function with scale parameter (the smoothing parameter) set at sm : $L(x, sm) = (1 + \exp(-sm \cdot x))^{-1}$. In addition we define $\bar{L}(x, sm) = 1 - L(x, sm)$. As the smoothing parameter sm increases the function $L(x, sm)$ becomes almost indistinguishable from the indicator function. Using the smoothed version of the indicator function (and other function we define below) is particularly important in the optimization step of the indirect inference approach.

The second set of auxiliary statistics captures the information required to calculate the bunching estimator i.e. the proportion of observations observed within approximately $100 \cdot q\%$ of the kink, adjusted for the difference in the net-of-tax rates, P_{kink} , and the average density of the earnings distribution within $100 \cdot r\%$ off the kink but not at the kink itself, $P_{\text{around kink}}$, with $r > q$.¹⁵

To define these statistics we use the smooth function $K(x, q, sm) = L(x + q, sm)\bar{L}(x - q, sm)/L(q, sm)^2$ to approximate the indicator $1_{|x| < q}$ for q positive.¹⁶ Furthermore, we set $\bar{K}(x, q, sm) = 1 - K(x, q, sm)$ which then approximates the indicator $1_{|x| > q}$.

In the specific context of our simulation experiment, where the unobserved components are normally distributed, and simplifying by assuming that the kink and the net-of-tax rates are constant over time, the quantity P_{kink} measures

$$\left(\Phi \left[\frac{1}{\sigma_\omega} (\ln k + q - \kappa - \alpha \ln \tau_1^c) \right] - \Phi \left[\frac{1}{\sigma_\omega} (\ln k - q - \kappa - \alpha \ln \tau_0^c) \right] \right) \frac{1}{\ln \frac{\tau_0^c}{\tau_1^c}}.$$

In the spirit of Saez's approximation, using the Taylor expansion to the first order we can approximate P_{kink} as

$$\phi \left[\frac{1}{\sigma_\omega} (\ln k - \kappa - \alpha \overline{\ln \tau^c}) \right] (2q + \alpha \ln \frac{\tau_0^c}{\tau_1^c}) \frac{1}{\ln \frac{\tau_0^c}{\tau_1^c}},$$

with $\overline{\ln \tau^c} = (\ln \tau_1^c + \ln \tau_0^c)/2$. Similarly, the density $P_{\text{around kink}}$ measures

$$\begin{aligned} & \left[\Phi \left[\frac{1}{\sigma_\omega} (\ln k + r - \kappa - \alpha \ln \tau_1^c) \right] - \Phi \left[\frac{1}{\sigma_\omega} (\ln k + q - \kappa - \alpha \ln \tau_1^c) \right] \right] \\ & + \Phi \left[\frac{1}{\sigma_\omega} (\ln k - q - \kappa - \alpha \ln \tau_0^c) \right] - \Phi \left[\frac{1}{\sigma_\omega} (\ln k - r - \kappa - \alpha \ln \tau_0^c) \right] \frac{1}{2(r - q)}, \end{aligned}$$

¹⁵Note that q , r and sm are fixed parameters as far as the estimation is concerned. In our Monte Carlo experiment we set $q = 0.001$, $r = 0.15$ and $sm = 100$.

¹⁶The adjustment by $L(q, sm)^2$ is only important if q is small, i.e. so that $q \cdot sm < 10$, and it ensures that $K(0, q, sm) = 1$. When $q \cdot sm > 10$, $L(q, sm)^2$ is approximately equal to 1.

which we can approximate as

$$\phi\left[\frac{1}{\sigma_\omega}(\ln k - \kappa - \alpha \overline{\ln \tau^c})\right].$$

The ratio of the two approximated quantities thus reduces to

$$\frac{P_{\text{kink}}}{P_{\text{around kink}}} \approx \frac{2q + \alpha \ln \frac{\tau_0^c}{\tau_1^c}}{\ln \frac{\tau_0^c}{\tau_1^c}}.$$

Hence for small q , we obtain an estimator $\tilde{\alpha}$ for α

$$\tilde{\alpha} \approx \frac{P_{\text{kink}}}{P_{\text{around kink}}}.$$

This expression shows that the ratio of the two auxiliary statistics is particularly informative about the value of the ETI. Furthermore, these two auxiliary statistics rely on level cross sectional information only.

The third set of auxiliary statistics captures the proportion of observations moving around the kink. These statistics are informative about the relative importance of the transitory term in the unobserved component relative to the ETI. In a world where the kink changes over time we expect the transitions around the kink to be unbalanced: the proportion of observations moving in one direction, i.e, from below to above the kink, may be larger than the proportion of observations moving in the other direction. We measure these transitions by considering the proportion of observations which are observed below the kink at time t and above the kink at time $t + 1$, $T_{\text{below/above}}$, and vice-versa, $T_{\text{above/below}}$.

Observe that $T_{\text{below/above}}$ is an estimate of the average over all periods of the probabilities

$$\Pr\left[\omega_t < z_t^{-k}, \omega_{t+1} > z_{t+1}^{+k}\right] = \Phi_2[z_t^{-k}/\sigma_\omega, -z_{t+1}^{+k}/\sigma_\omega, -\rho_{t,t+1}],$$

while $T_{\text{above/below}}$ measures the average of

$$\Pr\left[\omega_t > z_t^{+k}, \omega_{t+1} < z_{t+1}^{-k}\right] = \Phi_2[-z_t^{+k}/\sigma_\omega, z_{t+1}^{-k}/\sigma_\omega, -\rho_{t,t+1}],$$

where $z_t^{-k} = \ln k_t - \kappa - \alpha \ln \tau_{0t}^c$, $z_t^{+k} = \ln k_t - \kappa - \alpha \ln \tau_{1t}^c$ and $\rho_{t,t+1}$ is the correlation between ω_t and ω_{t+1} .

Consider now the difference between $T_{\text{below/above}}$ and $T_{\text{above/below}}$ at times t and $t + 1$ only. In general, this is

$$\begin{aligned} & T_{\text{below/above}} - T_{\text{above/below}} = \\ & \Pr\left[\omega_t < z_t^{-k}\right] - \Pr\left[\omega_{t+1} < z_{t+1}^{+k}\right] - \\ & \Pr\left[\omega_t < z_t^{-k}, \omega_{t+1} < z_{t+1}^{+k}\right] + \Pr\left[\omega_t < z_t^{+k}, \omega_{t+1} < z_{t+1}^{-k}\right]. \end{aligned}$$

Observe that if $z_t^{+k} = z_t^{-k}$ and $z_{t+1}^{+k} = z_{t+1}^{-k}$, then the previous expression simplifies to read

$$T_{\text{below/above}} - T_{\text{above/below}} = \Pr[\omega_t < z_t^{-k}] - \Pr[\omega_{t+1} < z_{t+1}^{+k}].$$

This suggests, therefore, to approximate the difference between the probabilities at the mean of the quantities z_t^{+k} and z_t^{-k} , and z_{t+1}^{+k} and z_{t+1}^{-k} , i.e.,

$$\begin{aligned} T_{\text{below/above}} - T_{\text{above/below}} &\approx \\ \Pr[\omega_t < (z_t^{-k} + z_t^{+k})/2] &- \Pr[\omega_{t+1} < (z_{t+1}^{-k} + z_{t+1}^{+k})/2]. \end{aligned}$$

To a first order in the log normal case this becomes

$$\begin{aligned} T_{\text{below/above}} - T_{\text{above/below}} &\approx \\ \frac{\alpha}{\sigma_\omega} (\overline{\ln \tau_t^c} - \overline{\ln \tau_{t+1}^c}) \phi &[\frac{1}{\sigma_\omega} (z_t^{-k} + z_t^{+k} + z_{t+1}^{-k} + z_{t+1}^{+k})/4], \end{aligned}$$

The first term in this approximation reflects the interaction between the ETI, α , and the variance of the unobserved component. The approximation suggest, therefore, that if the ETI is small, or if the average of the log of the net tax rates is constant over time, the difference between the two transition rates will be close to zero. Similarly if the overall variance, σ_ω^2 , is large, the two transition rates are nearly equal. Note that we can not consider the case when σ_ω^2 is small, since it would mean that there is no variability in the data either from a permanent source or a transitory source.

To account for the error structure, in the fourth set of statistics we measure the variance of log earnings around their year average, V_{level} , as well as the variance in earnings growth around their year average, V_{growth} . The first quantity captures the overall variability, while the second term only captures the variance of the transitory component.

Finally, we measure three covariances. The first two C_{level} and C_{growth} are auto-covariances, which are informative about the autocorrelation of the transitory component. The last covariance, $C_{\text{growth tax}}$ measures the association between the growth of the net-of-tax rate and earnings growth. We would expect this quantity to respond to the value of the ETI.

Means

$$M_{\text{level} < \text{kink}} = \frac{1}{NT} \sum_i \sum_{t=1} \ln wh_{it} \cdot L(\ln \frac{k_t}{wh_{it}}, sm),$$

$$M_{\text{level} > \text{kink}} = \frac{1}{NT} \sum_i \sum_{t=1} \ln wh_{it} \cdot \bar{L}(\ln \frac{k_t}{wh_{it}}, sm),$$

$$M_{\text{growth}} = \frac{1}{N(T-1)} \sum_i \sum_{t=2} \Delta \ln wh_{it},$$

Proportions

$$P_{\text{kink}} = \frac{1}{NT} \sum_i \sum_{t=1} K\left(\ln \frac{k_t}{wh_{it}}, q, sm\right) \frac{1}{\ln \frac{\tau_{0t}^c}{\tau_{1t}^c}},$$

$$P_{\text{around kink}} = \frac{1}{NT} \sum_i \sum_{t=1} K\left(\ln \frac{k_t}{wh_{it}}, r, sm\right) \bar{K}\left(\ln \frac{k_t}{wh_{it}}, q, sm\right)$$

Transitions

$$T_{\text{below/above}} = \frac{1}{N(T-1)} \sum_i \sum_{t=2} L\left(\ln \frac{k_{t-1}}{wh_{it-1}}, sm\right) \bar{L}\left(\ln \frac{k_t}{wh_{it}}, sm\right),$$

$$T_{\text{above/below}} = \frac{1}{N(T-1)} \sum_i \sum_{t=2} \bar{L}\left(\ln \frac{k_{t-1}}{wh_{it-1}}, sm\right) L\left(\ln \frac{k_t}{wh_{it}}, sm\right),$$

Variances

$$V_{\text{level}} = \frac{1}{NT} \sum_i \sum_t (\ln wh_{it} - \overline{\ln wh_{.t}})^2,$$

$$V_{\text{growth}} = \frac{1}{N(T-1)} \sum_i \sum_{t=2} (\Delta \ln wh_{it} - \overline{\Delta \ln wh_{.t}})^2,$$

Covariances

$$C_{\text{level}} = \frac{1}{N(T-1)} \sum_i \sum_{t=2} (\ln wh_{it-1} - \overline{\ln wh_{.t-1}})(\ln wh_{it} - \overline{\ln wh_{.t}}),$$

$$C_{\text{growth}} = \frac{1}{N(T-2)} \sum_i \sum_{t=3} (\Delta \ln wh_{it-1} - \overline{\Delta \ln wh_{.t-1}})(\Delta \ln wh_{it} - \overline{\Delta \ln wh_{.t}}),$$

$$C_{\text{growth tax}} = \frac{1}{N(T-1)} \sum_i \sum_{t=2} \Delta \ln wh_{it} \Delta \ln \tau_{it}^c.$$

We collect all these statistics into a vector s where

$$s = (M_{\text{level}<\text{kink}}, M_{\text{level}>\text{kink}}, M_{\text{growth}}, P_{\text{kink}}, P_{\text{around kink}}, T_{\text{below/above}}, T_{\text{above/below}}, V_{\text{level}}, V_{\text{growth}}, C_{\text{level}}, C_{\text{growth}}, C_{\text{growth tax}}).$$