NBER WORKING PAPER SERIES

# ADJUSTMENT COSTS, FIRM RESPONSES, AND LABOR SUPPLY ELASTICITIES: EVIDENCE FROM DANISH TAX RECORDS

Raj Chetty
John N. Friedman
Tore Olsen
Luigi Pistaferri

Working Paper 15617
http://www.nber.org/papers/w15617

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2009

We would like to thank David Card, Jim Hines, Henrik Kleven, Claus Kreiner, Patrick Kline, Robert Moffitt, John Pencavel, Emmanuel Saez, Esben Schultz, and numerous seminar participants for helpful suggestions and valuable discussion. We are extremely grateful to Mette Ejrnæs and Bertel Schjerning at the Centre for Applied Microeconometrics at the University of Copenhagen, Frederik Hansen at the Ministry of Finance, as well as Anders Frederiksen, Paul Bingley and Niels Chr. Westergård-Nielsen at Aarhus Business School for help with the data and institutional background. Gregory Bruich provided outstanding research assistance. Support for this research was provided by the Robert Wood Johnson Foundation and NSF Grant SES-0645396. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records
Raj Chetty, John N. Friedman, Tore Olsen, and Luigi Pistaferri
NBER Working Paper No. 15617
December 2009
JEL No. E62,H2,J22

## ABSTRACT

We show that adjustment costs for workers interact with hours constraints set by firms to determine the effect of taxes on labor supply. We present evidence supporting three predictions of an equilibrium model in which firms post wage-hours packages and workers pay search costs to find jobs. First, observed labor supply elasticities increase with the size of the tax variation from which they are identified. Second, tax changes that apply to a larger group of workers generate larger responses. Third, firms tailor job offers to match workers' tax preferences. Calibrating our model to match the empirical findings, we find that standard microeconometric methods underestimate structural labor supply elasticities by an order of magnitude.

Raj Chetty
Department of Economics
Harvard University
1805 Cambridge St.
Cambridge, MA 02138
and NBER
chetty@fas.harvard.edu

John N. Friedman
Harvard Kennedy School
Taubman 356
79 JFK St.
Cambridge, MA  02138
and NBER
jfriedman@post.harvard.edu

Tore Olsen
Department of Economics
Harvard University
1805 Cambridge St.
Cambridge, MA 02138
tolsen@fas.harvard.edu

Luigi Pistaferri
Department of Economics
579 Serra Mall
Stanford University
Stanford, CA 94305-6072
and NBER
pista@stanford.edu

# 1 Introduction

The vast literature on tax and transfer policies rests on the assumption that workers can freely choose jobs that suit their preferences (e.g. Mirrlees 1971). In this paper, we present quasi-experimental evidence that challenges this assumption. We show that the effect of taxes on labor supply is determined by two factors that limit workers' ability to make optimal choices: adjustment costs and hours constraints imposed by firms. Because of these factors, distortions in the jobs offered by firms shape the effects of tax policies on equilibrium labor supply. We show that these forces attenuate conventional micro-econometric estimates of labor supply elasticities, and develop techniques for recovering long-run elasticities from observed behavioral responses.

To motivate our empirical analysis, we develop a stylized model of labor supply and taxation with search costs and endogenous hours constraints. We consider a model where firms have constant-returns-to-scale Leontief production technologies and therefore have all of their employees work the same number of hours to exploit complementarities in production. Firms post hours offers, generating a distribution of wage-hours packages in the economy. Individuals draw job offers from the hours distribution for their wage level. If an individual is unsatisfied with the job she draws, she can search for a better match by paying a search cost. Because of search costs, individuals are distributed around their optimal level of hours. In equilibrium, the aggregate distribution of jobs offered by firms is determined by workers' preferences. This endogenous determination of wage-hours packages is the key innovation in this model relative to existing models in which firms' technologies fully determine the distribution of wage-hours packages (Rosen 1976, Blundell, Brewer, and Francesconi 2008).

The model generates three testable predictions about the relationship between taxes and labor supply. First, the observed elasticity of labor supply increases with the size of the tax variation from which the estimate is identified.[1] Intuitively, large tax changes prompt more individuals to respond, leading to a larger observed elasticity that more closely approximates the long-run effect of taxes. Analogously, the model predicts that individuals are more

[1] This prediction refers to tax changes starting from an economy that is initially in steady state. If the local elasticity varies with the tax rate, the testable prediction is that the observed response to one large tax change starting from steady state will exceed the sum of the observed response to two or more smaller tax changes (each starting from steady state) that add up to the total size of the large tax change.

responsive to larger kinks in the tax system, and the elasticities estimated from bunching at kink points therefore increase with the size of the kink. Second, the observed elasticity increases with the size of the group of workers affected by a tax change. Changes in taxes induce changes in labor supply not just by making individuals accept different job offers, but also by changing the distribution of hours offered by firms in equilibrium. Because changes in taxes that affect a larger group of individuals induce larger changes in hours constraints, they produce larger observed elasticities. Furthermore, a tax reform can affect even the labor supply of workers whose personal tax incentives are unchanged by distorting their coworkers' incentives and inducing changes in hours constraints. Finally, the model predicts a correlation between firm responses and individual responses to taxes. If firms cater to workers' tax preferences when offering jobs, then one should observe larger distortions in the equilibrium distribution of job offers in sectors where workers themselves exhibit larger tax elasticities.

We test these three predictions using an administrative matched employer-employee panel of the population in Denmark between 1994 and 2001. This dataset combines information on tax records, demographic characteristics, education, and employment characteristics such as occupation, tenure, and experience. There are two sources of tax variation in the data: tax reforms across years, which produce variation in marginal net-of-tax wage rates of less than 7%, and changes in tax rates across tax brackets within a year, which generate much larger variation in net wages, sometimes up to 30%. We focus primarily on the cross-bracket variation in taxes rates because it is larger and applies to the entire population, permitting coordinated responses. In particular, we estimate observed elasticities by studying the amount of bunching in earnings generated by kinks in the budget set, as in Saez (2009).

We first document that the degree of bunching at kink points varies substantially with the size of the kink, consistent with the first prediction of the model. We document substantial and visually evident excess mass near kink points at which the net-of-tax wage falls by more than 25%. We find modest excess mass at kinks where the net-of-tax wage falls by 10%, and no excess mass at kinks that generate variation in net-of-tax wages smaller than 10%. We find no changes in earnings around tax reforms that change net-of-tax wages by less than 10% across a wide range of initial tax rates. This finding shows that the increasing behavioral responses we detect are driven by adjustment costs rather than an elasticity of labor supply that varies with the tax rate. The observed elasticities at the largest kinks are nearly three times those observed at smaller kinks and implied by small tax reforms.

To test the second prediction, we exploit heterogeneity in the location of top tax bracket cutoffs across workers. Approximately 60% of wage earners have few deductions and little non-wage income, and therefore reach the top tax bracket when their wage earnings equal the top tax cutoff for taxable income. Workers with significant deductions or non-wage income, however, reach the top tax cutoff at different levels of wage earnings and thus have less common tax incentives. We first demonstrate that the mode of occupation-level wage earnings distributions has an excess propensity to be located near the top tax bracket cutoff, catering to modal worker preferences.[2] Importantly, even workers who have significant deductions or non-wage income, and therefore have no incentive to locate their wage earnings at the top tax cutoff, disproportionately gather at the top tax cutoff. This phenomenon, which we term "firm bunching," constitutes direct evidence that firms tailor wage-hours packages to match common tax preferences among workers. Next, we show that workers bunch less in occupations whose modal earnings level is far from the top tax cutoff, since these workers must actively search for an uncommon job to reach the kink. Finally, we find that firms synchronize wage earnings growth for workers near the kink with movement in the kink over time. Together, these results show that technological constraints that force workers to coordinate on similar hours and salaries have substantial impacts on behavioral responses to taxation.

Although firm bunching is an important source of behavioral responses to the tax system, some of the bunching we observe at kinks is driven by individual workers searching for jobs that place them near the top tax kink. To isolate and measure such "individual bunching," we exploit a cap on tax-deductible pension contributions, which is on average DKr 33,000 in the years we study. Approximately 1.5% of workers make pension contributions up to this amount and therefore cross into the highest income tax bracket when they earn DKr 33,000 more than the legislated top tax cutoff that applies for the 60% of the population that makes no pension contributions. We document that this pension-driven kink induces excess mass in the distribution of wage earnings at DKr 33,000 above the top tax cutoff. This excess mass is driven solely by individual responsess since there is no bunching at the pension-driven kink for workers who have small deductions. Hence, firms do not appear to cater to the preferences of

[2] We focus on wage earnings distributions at the occupation level because most workers' wages are set through collective bargains at the occupation level in Denmark.

this small group of workers. Instead, these workers must actively search for jobs that place them at the top tax kink given their pension contribution levels.

We test the third prediction by estimating the correlation between individual and firm bunching across occupations and demographic groups. We find that groups in which workers exhibit more individual bunching at the pension-driven kink are also those where firms are more likely to bunch workers at the legislated kinkpoint Although these results cannot be interpreted as causal effects because we do not have exogenous variation in the degree of individual bunching, they are consistent with the model's prediction that firms cater to workers' tax preferences in equilibrium. Consistent with the importance of firm responses, we find that some of the heterogeneity in observed elasticities across demographic groups is driven by occupational choice. For instance, reweighting men's occupations to match those of women's eliminates 70% of the gap in observed elasticities between men and women. We also find that self-employed individuals, who presumably have very low adjustment costs and are not subject to hours constraints imposed by firms, are an order of magnitude more likely to locate at kinks than wage earners.

Drawing upon these empirical findings, we show that reduced-form estimated elasticities in these data understate the long-run elasticity parameter by an order of magnitude. In our model, the structural elasticity is non-parametrically identified only as the size of the tax change grows arbitrarily large and the fraction of the population affected by the change goes to zero. Intuitively, the large size of such a tax change implies that all those affected respond, and so adjustment costs no longer matter; the small scope of such a tax change removes the interaction of individual choice with job offers from firms. Unfortunately, this identification-at-infinity approach cannot be implemented in practice because we must rely on finite tax changes. Instead, we show in a calibration of our model that an upper bound on the adjustment cost of 20% of consumption implies a lower bound of 0.28 on the structural elasticity. Intuitively, the observed elasticities must be attenuated at least to some degree due to the upward-sloping relationship observed in the data between the estimated elasticity and size of the underlying tax change. Quantitatively, we find that the size of adjustment costs required to attenuate smaller elasticities would be unreasonably high due to the relatively large utility losses from suboptimal labor supply choice. Therefore only larger elasticities are consistent both with reasonable adjustment costs and the patterns in our data.

Beyond demonstrating how adjustment costs and firm responses mediate the relationship between taxes and labor supply, the findings in this paper offer two broader lessons. First, they help reconcile the substantial difference between the labor supply elasticities estimated in existing microeconometric studies and macroeconomic studies. Micro studies generally find elasticities below 0.1 on the intensive-margin for all individuals excluding top income earners (Chetty 2009a). But macroeconomic studies find intensive-margin labor supply elasticities above 0.5 (e.g. Prescott 2004, Davis and Henrekson 2005, Rogerson and Wallenius 2008, Ohanian et al. 2008). The large, economy-wide, long-run variation exploited in macro studies may be more likely to overcome adjustment costs and induce changes in firm behavior in equilibrium, potentially explaining why they obtain larger elasticity estimates. Second, at a broader methodological level, the results call into question the modern paradigm of using small quasi-experiments that apply to narrow subgroups to learn about the effects of economic policies and shocks on behavior. In settings with rigid institutional structures and frictions in adjustment, the long-run effects of policies implemented at an economy-wide level are likely to be very different from the effects of quasi-experiments. Our analysis shows that long-run effects can be approximated by using very large, broad policy variation for identification coupled with a structural model.

Our study also builds on and contributes to three other strands of the literature. First, the literature on non-linear budget sets (e.g., Hausman 1981, Moffitt 1990, Saez 2002) generally finds no evidence of bunching. As noted by Blundell and MaCurdy (1999), "...for the vast majority of data sources currently used in the literature, only a trivial number of individuals, if indeed any at all, report [earnings] at interior kink points." The kinks examined in previous studies are generally much smaller – both in the change in tax rates at the kink and the size of the group of individuals affected – than the largest kinks studied here. Our calibration analysis shows that one can explain behavior both at and away from kinks by incorporating adjustment costs and hours constraints into a non-linear budget set model. Second, our empirical analysis relates to recent work on taxable income elasticities vs. labor supply elasticities. Feldstein (1995) shows that taxable income elasticities are a sufficient statistic for tax policy analysis, but more recent studies have argued that it is important to distinguish income shifting from "real" changes in earnings choices (Moffitt and Wilhelm 2000, Goolsbee 2000, Chetty 2009b). Because our dataset has several measures of compensation, including taxable income, nontaxable compensation, and wage earnings, we are able to distinguish between income shifting and "real" labor supply responses more precisely than earlier work. In particular, we show that the excess mass observed at large kinks is driven by wage earnings and is not due to shifting into untaxed forms of income such as pension contributions. Finally, our findings provide evidence in support of a theoretical literature which has argued that the effects of government policies may operate primarily through coordinated changes in social norms or institutions rather than individual behavior (e.g. Lindbeck 1995, Alesina, Glaeser and Sacerdote 2005).

The paper is organized as follows. In Section 2, we set up a stylized model of adjustment costs and labor supply and derive testable predictions. Section 3 describes the Danish data and provides some institutional background. Section 4 presents the empirical results. Section 5 presents a calibration of our model to produce an estimate of the of underlying elasticity parameter implies by the data, and section 6 concludes.

## 2 Labor Supply and Taxes in a Model with Frictions

We motivate our empirical analysis using a stylized model of labor supply with search costs and endogenous hours constraints. We show how observed labor supply elasticities vary with the size and scope of the tax changes used for identification. In our empirical analysis, we exploit variation in taxes across brackets by examining bunching at kinks and variation in tax rates from legislated tax reforms. We therefore consider both sources of variation in the model.

### 2.1 Setup

Consider a model in which workers live for two periods, after which they are replaced by a new (non-overlapping) generation.[3] Individuals, indexed by $i$, have quasi-linear flow utility

$$u_i\left(c_i, h_i\right) = c_i - \alpha_i^{-\frac{1}{\varepsilon}} \frac{h_i^{1+\frac{1}{\varepsilon}}}{1+\frac{1}{\varepsilon}} \tag{1}$$

[3] Though this model is essentially static, we believe that our results would hold in a life-cycle model with dynamic concerns.

over a numeraire consumption good $c_i$ and hours of work $h_i$.[4] This quasi-linear utility function is analytically convenient because it eliminates income effects and intertemporal substitution. We abstract from income effects because we focus primarily on local variation in tax rates at kink points in our empirical analysis, which generates negligible income effects. The parameter $\alpha_i$ captures heterogeneity in disutilities of labor supply. Let $\phi_i$ denote the "adjustment cost" for each worker, a parameter whose importance will become clear below. Disutility $\alpha_i$ and adjustment cost $\phi_i$ is drawn from a smooth distribution $F(\cdot,\cdot)$ with full support on a compact subset of the non-negative real numbers. Workers all have the same marginal product. Let $w_i$ denote the wage rate paid to worker $i$, determined by firms in equilibrium.

Workers face a progressive two-bracket income tax system. To characterize tax reforms that affect subgroups of workers, we assume that there are two types of tax systems, $s = \{1, 2\}$. A fraction $\zeta$ of workers face tax system 1 and the remainder $(1-\zeta)$ face tax system 2. Individual of type $s = 1$ face a tax system with two brackets; the tax rate in the lower bracket is $\tau_1$ and the tax rate in the upper bracket is $\tau_2 > \tau_1$, and these workers begin to pay the higher tax rate when their earnings exceed $K$, a kink point. However, those of type $s = 2$ pay the lower rate $\tau_1$ on all of their income. An individual $i$ in tax group $s$ has consumption

$$\begin{aligned} c_i(h_i|s=1) &= (1-\tau_1)\min(wh_i, K) + (1-\tau_2)\max(wh_i - K, 0) \\ c_i(h_i|s=2) &= (1-\tau_1)\,wh_i. \end{aligned}$$

We omit unearned income because the quasi-linear functional form (1) implies no income effects in labor supply. Denote a worker's optimal hours choice by $h^*(\alpha, s)$.

Firms of fixed size $N$ employ workers to produce goods sold for price $p$. Firms have a constant returns to scale Leontief production function in terms of the hours worked by workers at the firm. Firm profits are

$$\pi_j = pN \min_{i\in j}\{h_i\} - w_j \sum_{i\in j} h_i$$

where $i \in j$ denotes that worker $i$ works at firm $j$. Given this technology, a firm offers a single

[4]This utility specification implies a constant wage elasticity of labor supply. We extend our analysis to utility functions that generate non-constant elasticities below.

level of hours $h_i = h_j$ for all of its workers. For instance, factory production requires that all workers work for the same shift length, effectively generating hours constraints. Firms thus offer workers a contract specifying $h_j$ and $w_j$.

We model labor market frictions using the following job search process. First, firms post a wage-hours package. Let the distribution of hours offered be denoted by $G(h)$. Workers then search for jobs and draw one randomly from the distribution of jobs offered by firms.[5] Each worker $i$ may then pay to search narrowly for a better job. In this case, she will draw a job offering her optimal hours with probability $q$, or a random job from a diffuse distribution $G'(h'|h^*)$ centered around the optimal job with probability $1-q$. This probability $q$ is endogenous; the cost of choosing $q$ is $\phi_i c(q)$ for some function $c(\cdot) \geq 0$ where $c'(\cdot) > 0$.[6] Thus, workers paying to find a new job may optimize the precision of their search. We normalize the $c(q=1) = 1$, so that the parameter $\phi_i$ may be interpreted as the cost of choosing one's optimal job perfectly. We also assume that workers can costlessly switch between firms required a given hours amount $h$, and so each such "slice" of the labor market is competitive. If the tax system does not change in between periods, then of course there is no need for reoptimization. But if there is a reform, then firm may costlessly change the hours-wage package, after which each worker may once again opt to pay $\phi_i c(q)$ to find a firm offering her optimal labor supply.

In equilibrium, workers maximize utility given search costs and wages, firms maximize profits by choosing the size of the workforce $l$, and markets clear, so that the number of jobs supplied with a given hours level $h$ equals the number of workers supplying that level $h$ at wage $w$. Since firms maximize profits in a competitive "slice," we know that $w_j = p$ for all $i$, since workers must be paid their marginal products. At every level of hours $h$, the fraction of workers who demand that job after drawing and endogenously opting in or out of costly search must equal the fraction of jobs offered at that hours level $h$ to begin with. In other words, equilibrium involves a distribution of job offers $G(h)$ that is a fixed point of the functional that is the job-search process.

We define two horizons over which tax policies affect the economy. The "short run" refers

[5] Since we are already looking only at the slice of the labor market where workers have productivity $w$, this assumption is not particularly extreme. Nevertheless, a somewhat directed initial search does not alter the main results below so long as the follow-up search is more precise.

[6] Note that we place no restriction on $c(0) \geq 0$, so that workers may pay a fixed cost to begin to search at all.

to the effects of tax changes within a generation of workers. This horizon holds the set of workers in the economy constant, but allows workers to search for a new job if they so choose, and also allows firms to reoptimize their hours constraints. In the "long run," old workers die and are replaced by a new generation that optimizes with respect to the tax system in place when they are born. Because these new workers are not bound by the choices of their predecessors, tax policies may have different effects in the short run and long run.

This general equilibrium model with search costs is analytically intractable. To illustrate the key predictions and intuitions of the model, we first focus on two special cases that we can analyze analytically. First, we characterize the impact of taxes on worker behavior in partial equilibrium, holding the distribution of firm offers fixed. We then consider the impact of firm optimization in general equilibrium when the distribution of adjustment costs for workers is degenerate: $\phi_i \in \{0, \infty\}$. Finally, we present numerical simulations showing that the results proved in the special cases hold under plausible parametrizations of the general model.

## 2.2 Benchmark: Effects of Taxes without Frictions

We begin by reviewing the effects of taxes on labor supply in the standard model used in the income tax literature, where there are no frictions. When $\phi_i = 0$ for all workers, each worker may freely choose to search for a job with $q = 1$, and so labor supply is determined purely by worker choice, with $h_i = h_i^* (\alpha_i; s)$ for all workers.

The literature focuses on identifying the structural parameter $\varepsilon$ because it fully determines the effects of changes in taxes and wages on labor supply in the frictionless model. To see this, consider the effects of an unanticipated increase in $\tau_1$ between periods 1 and 2 of a generation that affects only workers facing the tax schedule $s = 1$. To focus purely on the effects of the change in marginal rates, we focus on the special case of our model where $K \to \infty$, and thus the tax system contains only one bracket with rate $\tau_1$. Workers with $s = 1$ reduce their hours $h$ such that

$$\frac{d \log h}{d \log(1 - \tau_1)} = \varepsilon. \tag{2}$$

Those with $s = 2$ , who are unaffected by the reform, do not change hours of work.[7] Hence,

[7] For example, workers of type $s = 1$ could be married and workers of type $s = 2$ single taxpayers. In this example, the tax reform we are considering consists of an increase in the marginal tax rate faced by married taxpayers (the introduction of a "marriage penalty").

in this frictionless model, the observed elasticity of hours with respect to the net-of-tax rate $(1-\tau_1)$ for those affected by the tax reform coincides with the structural parameter of interest. We shall therefore refer to $\varepsilon$ as the "structural" elasticity. Note that estimate $\varepsilon$ depends on neither the size of the reform nor the fraction $\zeta$ of workers affected by it.

The elasticity $\varepsilon$ is most commonly estimated using variation in tax rates from tax reforms (see Blundell and MaCurdy 1999, Saez et al. 2009). However, $\varepsilon$ can also be identified from cross-sectional variation using the amount of bunching observed at the kink, as shown by Saez (2009). Denote the fraction of individuals who locate at the kink by setting $h_i = \frac{K}{w}$ by $B\left(\varepsilon, \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)\right)$. Under the assumption that the distribution of earnings is locally log-linear, Saez (2009) shows that

$$\varepsilon \simeq \frac{B}{\ln\left(\frac{1-\tau_1}{1-\tau_2}\right)}. \tag{3}$$

This equation shows that in a frictionless model, one can identify the elasticity of labor supply in the cross section without any restrictions on the distribution of heterogeneity $\alpha_i$ by measuring the excess mass at kinks in the tax system.

An important property of equations (2) and (3) is that the observed elasticity does not depend upon the size of the tax change, nor upon the number of workers $\zeta$ affected by the tax change. Irrespective of the size and scope variation that is used, the observed elasticity always coincides with the structural elasticity in the standard frictionless model.

### 2.3 Special Case 1: Fixed Firm Offer Distribution

To characterize behavior in the presence of frictions, we first analyze the choice of workers who are initially matched to a firm offering hours $h_0$. With the non-linear budget set created by the tax system, a worker's optimal hours choice is

$$h_i^*(\alpha_i; s) = \begin{cases} \alpha_i (w_i (1-\tau_1))^{\varepsilon} & \text{if } \alpha_i < \underline{\alpha} \text{ and } s = 1 \text{ or if } s = 2 \\ h_K = \frac{K}{w_i} & \text{if } \alpha_i \in [\underline{\alpha}, \overline{\alpha}] \text{ and } s = 1 \\ \alpha_i (w_i (1-\tau_2))^{\varepsilon} & \text{if } \alpha_i > \overline{\alpha} \text{ and } s = 1 \end{cases}. \tag{4}$$

where $\underline{\alpha} = \frac{K}{(w_i(1-\tau_1))^{\varepsilon}}$ and $\overline{\alpha}_s = \frac{K}{(w_i(1-\tau_2))^{\varepsilon}}$. Intuitively, workers with low disutilities of labor (high $\alpha_i$'s) earn enough to reach the interior of the top bracket and choose labor supply based on $\tau_2$; workers with high disutilities of labor are in the interior of the lower bracket and choose based on $\tau_1$; and the set of workers with moderate disutilities "bunch" at the kink because the net-of-tax wage drops discontinuously at $K$.

A worker of type $\alpha_i$ facing tax schedule $s$ chooses her optimal search pattern in two steps. First, she chooses the precision of her search $q$. Since a worker's initial job draw has no influence on the secondary search process, all workers of type $\{\alpha_i, s\}$ would choose the same precision $q^*(\alpha_i, s)$ *if* they were to search. Second, given $q^*(\alpha_i, s)$, a worker chooses to search for a better job or not. This initial decision does depend on her initial job $h_0$; she searches for a new job if $h_0 \notin \left[\underline{h}_i, \overline{h}_i\right]$, where these thresholds are defined by the two equations:

$$Eu|_{q^*(\alpha_i,s)} - u\left(c_i(\underline{h}_i), \underline{h}_i\right) = \frac{\phi_i c\left(q^*\left(\alpha_i, s\right)\right)}{2} \text{ for } \underline{h}_i < h_i^* \tag{5}$$

$$Eu|_{q^*(\alpha_i,s)} - u\left(c_i(\overline{h}_i), \overline{h}_i\right) = \frac{\phi_i c\left(q^*\left(\alpha_i, s\right)\right)}{2} \text{ for } \overline{h}_i > h_i^* \tag{6}$$

where $Eu|_{q^*(\alpha_i,s)}$ denotes the expected utility derived from optimal search. Those drawing hours packages that fall within this threshold, so that $h_0 \in \left[\underline{h}_i, \overline{h}_i\right]$, stay at their initial job draw. Thus, after the search process is complete, there are two types of workers at each firm: a point-mass whose optimal labor supply $h^*$ is exactly that offered by the firm and a diffuse distribution of workers with optimal hours near but not equal to $h_0$.

We now show how worker frictions affect the mapping from observed behavioral responses to the structural elasticity in equation (3). In the model without search frictions, workers locate at the kink if and only if $\alpha_i \in [\underline{\alpha}, \bar{\alpha}]$ and $s = 1$. Denote the fraction of such workers in the population by $B\left(\varepsilon, \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)\right) = \zeta\rho = \zeta\left(F\left(\bar{\alpha}\right) - F\left(\underline{\alpha}\right)\right)$.

There are two conceptually distinct sources of bunching at the kink in the presence of frictions. The first, which we label "individual bunching," arises from workers whose ideal hours are $h_i^* = h_K$ and who choose to pay the search cost $\phi_i$ to find a job that places them there. Denote by $\lambda = \int\int\left[1 - G\left(\overline{h}_i\left(\alpha\right)\right) - G\left(\underline{h}_i\left(\alpha\right)\right)\right] dF\left(a, \phi \mid \alpha_i \in [\underline{\alpha}, \overline{\alpha}]\right)$ the fraction of such workers who choose to search for a better job. Thus, the total amount of individual bunching $B_I = \lambda\zeta\rho$. In the frictionless case, $\lambda = 1$, and so individual bunching accounts for all of the excess mass at the kink.

The second source of bunching at kinks, which we label "firm bunching," arises from workers who happen to get matched to firms that offer wage-hours packages that place workers at the kink. If the distribution of job offers is smooth, then a measure zero of workers fall into this category; but if the distribution of job offers includes itself a point-mass at the kink, which in some cases it may, then a positive mass of workers draw $h_0 = h_K$. Some of these workers, by chance, actually do want to locate at the kink, but others have ideal hours $h_i^*$ that differ from $h_K$. These workers choose to remain at this individually inefficient point because the utility loss from doing so is less than $\phi_i c(q_i^*)$. Denote by $\rho_s^+$ the share of workers with $\alpha_i \in \left[\underline{\alpha}_{si}^+, \bar{\alpha}_{si}^+\right]$, where

$$\begin{aligned} Eu|_{q^*\left(\underline{\alpha}_{si}^+, s\right)} - u\left(c_i(h_K), h_K\right) &= \frac{\phi_i c\left(q^*\left(\underline{\alpha}_{si}^+, s\right)\right)}{2} \text{ for } \underline{\alpha}_{si}^+ < \underline{\alpha} && (7) \\ Eu|_{q^*\left(\bar{\alpha}_{si}^+, s\right)} - u\left(c_i(h_K), h_K\right) &= \frac{\phi_i c\left(q^*\left(\bar{\alpha}_{si}^+, s\right)\right)}{2} \text{ for } \bar{\alpha}_{si}^+ > \bar{\alpha}. && (8) \end{aligned}$$

for $s = 1, 2$. (Note here that *both* workers of type $s = 1$ who face the kink *and* $s = 2$ who do not face the kink may end up near the kink through "firm bunching.") Thus, denote the total amount of firm bunching $B_{F,s}$. The signature of firm bunching is that it induces bunching at common kink points even among workers whose optimal level of labor supply is *not* at the kink.[8]

Combining these equations shows that the total amount of bunching observed at kink $s$ is

$$\widehat{B} = B_{F,1} + B_{F,2} + B_I = \theta B\left(\varepsilon, \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)\right)$$

Let $\hat{\varepsilon}_s\left(\ln\frac{1-\tau_1}{1-\tau_2}, B_s\right)$ denote the elasticity estimated from inverting equation (3), which we shall refer to as the "observed" or "measured" elasticity. With search costs, the observed elasticity may not equal the structural elasticity, since

$$\hat{\varepsilon} = \theta\varepsilon. \tag{9}$$

We now characterize how the various parameters of our model affect $\theta$ and the observed elasticity.

[8] Note that this group also contains some workers who randomly draw $h_0 = h^* = h_{K_s}$. Though these workers are at their personal optimum, we include them as firm bunchers because they do not pay the search cost.

Proposition 1: *Without frictions, the observed elasticity is constant with respect to changes in the size of the tax change:*

$$\left.\frac{\partial \hat{\varepsilon}_1}{\partial \ln\left(\frac{1-\tau_1'}{1-\tau_1}\right)}\right|_{\phi=0} = 0$$

*but with frictions, the observed elasticity rises with both the size and scope of the tax change:*

$$\left.\frac{\partial \hat{\varepsilon}_1}{\partial \ln\left(\frac{1-\tau_1'}{1-\tau_1}\right)}\right|_{\phi>0} = 0.$$

Since we have held the firm offer distribution constant, this effect comes entirely through individual bunching. Intuitively, workers with $h^* = h_K$ *both* search with greater precision *and* pay to search more frequently when the change in incentives at the kink is larger. As $(\tau_2 - \tau_1)$ increases, first the optimal search precision increases:

$$\frac{\partial q^*}{\partial \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)} > 0.$$

Second, the region of inaction $\left[\underline{h}_i, \overline{h}_i\right]$ where a worker with $\alpha_i \in [\underline{\alpha}, \bar{\alpha}]$ accepts her initial hours offer becomes narrower:

$$\frac{\partial \underline{h}_i}{\partial \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)} > 0 \quad \text{and} \quad \frac{\partial \overline{h}_i}{\partial \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)} < 0. \tag{10}$$

These two effects each increase the amount of bunching, since a larger fraction of workers who wish to locate at the kink search, and then a larger fraction of these searching workers draw the kink exactly.

Figure 1 illustrates the intuition for this effect using indifference curves in consumption-labor space for a worker who would optimally locate at the kink. The bounds $\left[\underline{h}_i, \overline{h}_i\right]$ are where the budget constraint crosses the indifference curve that yields utility $\phi$ units less than the maximal utility $U^*$.[9] Now suppose $\tau_2$ increases, moving the upper budget segment from that with slope $(1-\tau_2)\,w$ to the one with slope $(1-\tau_2')\,w$. Then the utility loss from supplying

[9] For graphical simplicity, this is the indifference curve is $q = 1$; it is clear that a similar effect would hold for other indifference curves.

hours above the kink rises, as one earns less for this extra effort. As a result, the upper bound $\overline{h}_i$ decreases, which increases $\lambda$ and therefore $\theta$. Second,.the worker searches with greater precision since the cost of drawing from a diffuse distribution with probability $1-q$ becomes more costly.

Note that firm bunching may lead to $\theta > 1$, especially for large kinks. As the kink gets very large, most of the workers who should optimally locate at the kink actually do locate there, either through their initial match or by paying the search cost. Because many firms will offer jobs that place individuals at the kink, there will also be many suboptimal firm bunchers who locate at the kink. Adding those who optimally relocate to the kink to those who remain there suboptimally after the initial draw, the total mass at the kink can exceed the fraction of workers with $h_i^* = h_K$.

A second result is that, as the search costs decrease (so as $\phi_i$ decreases), the bounds of non-response $\left[\underline{h}_i, \overline{h}_i\right]$ shrink, implying $\frac{\partial \lambda}{\partial \phi} < 0$. Furthermore, the optimal precision of the search increases, as workers find it worthwhile to pay for a more finely tuned search. For individual bunchers, a decrease in the search cost has a similar effect to an increase in the size of the tax kink. Individual bunching $B_I$ is greater for populations with smaller search costs, as is the observed elasticity. However, the amount of firm bunching decreases as more people shift away from an initial job at the tax kink, so the impact of changes in $\phi$ on the total amount of bunching is indeterminate. Therefore we do not focus on this result in the empirical results below.

### 2.4 Special Case 2: Degenerate Search Cost Distribution

We next characterize the behavior of firms as they set the equilibrium job offer distribution. To obtain analytical results, we focus on a special case of the general model in which all workers have either $\phi_i = 0$ or $\phi_i = \infty$. Suppose that fraction $\delta$ of workers have $\phi_i = 0$. Furthermore, suppose that the distribution of $\alpha_i$, the disutility of labor, is independent of the individual adjustment cost.

The equilibrium distribution of job offers here must be the distribution of optimal hours choices $h_i^*$. Intuitively, those workers with $\phi_i = 0$ always choose their optimal point, and so the fixed point distribution for them must be the distribution of $h^*$. The two groups must have the same offer distribution, though, and any distribution would suffice as a fixed-point distribution for the $\phi_i = \infty$ group. Therefore, the total amount of bunching, in this special case, always equals the amount of bunching in a frictionless model. There is a key difference, though: in this model, fraction $1 - \delta$ of workers who bunch at kink $K$ have not chosen to be there, and thus are "firm bunchers." What is more, many of those workers may not even see a kink in their budget set at income $K$, since they are type $s = 2$.

Consider then those workers who locate, in equilibrium, at the kink $K$. Fraction $\zeta\rho$ of workers work hours $h_K$. But only fraction $\delta$ those workers have finite adjustment costs, and therefore we can write the total amount of individual bunching as $B_I = \delta\zeta\rho$. By definition, the remaining $(1-\delta)\,\zeta\rho$ fraction of workers with $h_K$ did not relocate there, and so $B_F = (1-\delta)\,\zeta\rho$. We can also calculate the equilibrium amount of bunching from each type of worker $B_s$. Manipulating the above formulas implies that, for workers who face tax schedule $s = 1$ and the tax kink,

$$B_1 = B_I + B_{F,1} = \delta\zeta\rho + (1-\delta)\,\zeta^2\rho.$$

The two terms in this expression represent individual bunching and firm bunching, respectively. Applying the formula from Saez (2009) in this case implies that

$$\hat{\varepsilon}_1 = \frac{B_1}{\zeta} = \delta\zeta\rho + (1-\delta)\,\zeta^2\rho. \tag{11}$$

The accounting identity $B = B_1 + B_2$ implies that

$$B_2 = B_{F,2} = (1-\delta)\,(1-\zeta)\,\zeta\rho.$$

This leads to our second comparative static:.

Proposition 2: *Without frictions, the observed elasticity is constant with respect to changes in the size of the tax change:*

$$\left.\frac{\partial\hat{\varepsilon}_1}{\partial\zeta}\right|_{\delta=0} = 0$$

*but with frictions, the observed elasticity rises with both the size and scope of the tax change:*

$$\left.\frac{\partial\hat{\varepsilon}_1}{\partial\zeta}\right|_{\delta>0} = 0.$$

We can easily see the effect by noting that the total mass of workers who face kinkpoint $K$ and who bunch at $h_K$ is the sum of $\delta\zeta\rho$, the workers who choose to relocate there, and $(1-\delta)\,\zeta^2\rho$ total mass of workers with infinite adjustment costs who draw $h_K$ initially (and still face the kink in the tax schedule). This second group of workers constitutes firm bunching, since they themselves may not want to locate at the kink, but due to the adjustment costs they accept their initial offer. Because of the firm bunchers, the fraction of worker bunching at the kink increases faster than proportionally as $\zeta$ increases, and so the estimated elasticity increases. Intuitively, the number of firm bunchers is larger at kink points faced by many people, because firms tailor jobs to match common preferences. At kinks faced by only a few people, very few workers draw $h_0 = h_K$, and virtually all of the mass must accumulate through individual bunching. But when $\zeta$ is large, many individuals locate at the kink via firm bunching, which increases the observed elasticity.

In order to investigate this prediction empirically, it will be helpful to develop empirically independent measures of firm bunching and individual bunching. We know from above that $B_I = \delta\zeta\rho$ and $B_F = (1-\delta)\,\zeta\rho$. We cannot directly use these formulas, because we observe neither $\rho$ nor $\delta$ in the data. We do however observe the tax schedule $s$ faced by each worker, and therefore we can independently estimate the amount of bunching $\hat{B}_s$ for each type of worker. The formulas above then imply that a straightforward empirical measure of firm bunching is

$$\hat{B}_F = \frac{\hat{B}_2}{1-\zeta}.$$

Intuitively, workers *cannot* optimally wish to bunch at a tax kink from someone else's tax schedule. As a result, *any* bunching among those not facing the tax kink is direct evidence of firm bunching.

In order to *independently* estimate individual bunching, we must make use of $\hat{B}_1$. In most settings, though, the estimate $\hat{B}_1$ includes both firm and individual bunching. As the scope of the tax kink converges to 0, though, the term $B_{F,1}$ converges to 0 at faster rate than $B_I$. Mathematically, it is easy to show that

$$\lim_{\zeta\to 0} \frac{\hat{B}_1}{B_I} = 1$$

Therefore, we can estimate the amount of firm bunching as

$$\hat{B}_I = \lim_{\zeta \to 0} \hat{B}_1. \tag{12}$$

To see the intuition behind this formula, recall the intuition that firms tailor the distribution of job offers to common worker preferences. As the share of workers facing the kink in tax shrinks to zero, firms have a decreasing incentive to offer jobs located at the kink. In the limit, when only measure zero of workers face the kink, the point mass at the kink in the distribution of jobs disappears. Workers then have a measure zero chance of initially drawing the kink, and so firm bunching disappears entirely. What remains in $\hat{B}_1$ is a pure measure of individual bunching.

With these measures in hand, our third and final prediction relates the amounts of firm and individual bunching:

PROPOSITION 3: *The amount of firm bunching and individual bunching should positively covary across different groups $g$ of the population with different elasticities:*

$$Cov\left(B_{I,g}, B_{F,g}\right)\big|_{\phi_i} > 0$$

Recall from above that the amount of individual bunching is $\delta\zeta\rho$ and that the amount of firm bunching is $(1-\delta)\,\zeta\rho$. Holding fixed the fraction $(1-\delta)$ of workers who cannot adjust, the fraction of workers bunching through each of these channels is fixed. Therefore as individual bunching increases or decreases, so too will the amount of firm bunching. As the structural elasticity parameter $\varepsilon$ increases, the amount of bunching in the frictionless case $\zeta\rho$, increases. Since both firm bunching and individual bunching are proportional to the amount frictionless bunching, we should therefore expect these two different forms of bunching to positively covary across groups in the population, to this extent that the groups differ in their structural elasticity parameter and not the cost of search.

## 2.5 General Case: Numerical Results

We now show that the results we derived in the special cases above also hold in the general model for plausible parameter values.[10] We assume that the cost function of precision is $c(q) = 1 + q^{1+\gamma}$. We then choose parameters that correspond to those used in the calibration of the model to the empirical evidence in Section 5. We choose the distribution of tastes to match the empirical distribution of income around the top kink in the Danish data. We assume that the diffuse non-directed search follows a normal distribution with standard deviation matching the year-on-year change in income, which we calibrate from the data as DKr $40,000$. We solve for the equilibrium hours offer distribution by iteration (a la Blackwell's theorem), and then check that we have found the unique equilibrium by starting the iteration from different points. To approximate non-wage income in the data, we add a random draw from a normal distribution with SD equal to DKr $5,000$ to each worker's labor income. As in the empirical work that follows, we estimate the counterfactual polynomial as a seven-degree polynomial on the equilibrium distribution within DKr $50,000$ on either side of the tax kink.

Our first prediction is that the estimated elasticity increases with the size of the tax kink. Figure 2a plots the estimated elasticity on the vertical axis for a number of differently sized tax kinks on the horizontal axis. As the log change in the net-of-tax rate increases from 3% to 30%, the estimated elasticity rises from 0.003 to 0.065. Note that the structural elasticity parameter $\varepsilon = 0.4$, so that the observed elasticity is still significantly less than the true value even at a relatively large kink. Our second prediction is that the estimated elasticity increases with the fraction of the population that shares the kink. Figure 2b plots the estimated elasticity on the vertical axis as the fraction of workers who face the kink ($\zeta$) varies on the horizontal axis. As the share of workers who face the kink at point $K$ increases from 10% to 100%, the estimated elasticity increases from less than 0.03 to just above 0.19 at 100%. The third prediction is that the amount of firm bunching and individual bunching positively covary as one changes the structural elasticity parameter. Figure 2c shows a scatterplot of the amount of firm bunching (vertical axis) and individual bunching (horizontal axis) for populations with different elasticities, holding the size of the kink and the scale of the adjustment costs fixed. Unlike in the stylized model above, the amount of individual bunching and firm bunching are

[10] See Appendix B for more detail on the structure of the numerical model.

no longer perfectly proportional. But the amount of individual bunching and firm bunching still move together quite closely in the model across the different groups.

### 2.6 Other Theoretical Results

*Distinguishing the Effects of Frictions from Non-Constant Elasticities.* An important simplifying assumption we made above is that the structural elasticity $\varepsilon$ is constant. In a model where the utility is not isoelastic, one may observe an elasticity that increases with the size of the tax change even without frictions. We can distinguish these two models by comparing the effects of several small tax changes with the effects of a larger change that spans the smaller changes. A model without adjustment costs predicts that small changes, each starting from steady state, should produce the same response as a single larger change. In contrast, the adjustment cost model predicts a larger response to the single large tax change. The same logic holds for bunching at kinks: the total amount of bunching at two small kinks is less than the bunching observed at one large kink in a model with adjustment costs, but not in a model with non-constant utility.

We establish this result formally in the appendix by analyzing a model with arbitrary utility $u(c, l)$. We show that for any tax rates $\tau_1 < \tau_2 < \tau_3$, a model without adjustment costs predicts the relationship

$$B(\tau_1, \tau_3) = B(\tau_1, \tau_2) + B(\tau_2, \tau_3).$$

The amount of bunching created from two smaller kinks is exactly equal to the bunching created at one larger kink. This is because the amount of bunching increases linearly with the size of the kink without adjustment costs, as shown in (3). The adjustment costs model instead predicts

$$\hat{B}(\tau_1, \tau_3) > \hat{B}(\tau_1, \tau_2) + \hat{B}(\tau_2, \tau_3).$$

Since the amount of bunching increases faster than proportionately with the size of the tax kink, the two smaller kinks produce less bunching than one larger kink. It is straightforward to establish a similar result for tax reforms.[11]

[11] When analyzing tax reforms, it is important that each small tax reform is levied on a different generation of workers in order to ensure that the population begins from steady state. If the same generation of workers face

*Identification Using Tax Reforms.* Our three results above analyze the effects of various parameters on the observed tax elasticity when estimated through bunching at kink points. Most studies in the empirical tax literature instead use tax reforms for identification; we now show that our results apply in this case too.

To characterize the impact of search costs on equation (2) most transparently, we consider the special case where $K = 0$ so that all workers initially face a linear tax schedule with rate $\tau$. Suppose the government increases the tax rate to $\tau' > \tau$ at the beginning of period 2 for those facing tax schedule $s = 1$. Workers do not anticipate this reform, and therefore make the same hours choices as above in period 1. Once the reform is announced, firms may change the hours worked at their firms, after which each worker can either stay at the same firm or reoptimize their hours by paying the search cost $\phi c(q)$.

The tax change affects labor supply through two channels. First, firms may change the wage-hours package offered in response to the new labor supply curve, and the labor supply of workers will follow. Recall from above that the equilibrium distribution of jobs matches the distribution of worker optimal jobs. For an worker of type $\alpha_i$, optimal labor supply is now

$$h_i^* = \alpha_i \left(p\left(1 - \tau'\right)\right)^{\varepsilon}$$

if they face schedule $s = 1$ and unchanged if they face $s = 2$.

The second channel for response to a tax change is that workers may choose to pay to reoptimize and search for a different job. This reoptimization can itself be of two types. First, those workers affected by the tax change may choose to search for a different job if they work for an unaffected entrepreneur. This channel is most important when $\zeta$ is small, that is for tax changes affecting a very small fraction of workers, who are unlikely to work for a firm whose entrepreneur shares their tax preferences. Second, workers *not* affected by the reform may search to undo the changes made by firms. These changes are the converse of the previous category; thus, this latter type of change is most likely when $\zeta$ is high and the tax change affects most but not all of the population. In the general model, the magnitude of this second effect can vary with a number of parameters. Therefore, as before, we derive

multiple unanticipated tax reforms, workers may pile up at the edges of their bounds, creating a large response to even a small further tax increase.

the two main predictions of our model in the two special cases.

In our first special case, we analyzed the behavior of workers in partial equilibrium. After experiencing a tax reform, workers' bounds of non-adjustment $\left[\underline{h}_i, \bar{h}_i\right]$ shift along with their optimal labor supply $h_i^*$. Workers will only now pay if their old job falls outside the new bounds. Since a larger tax reform shifts the optimal labor supply, and therefore the bounds, by more, it must therefore be that workers are more likely to pay to switch in response to a larger tax change. Intuitively, it does not pay to respond to a small tax change. This generates the first empirical prediction.

We develop the second and third empirical predictions in the second special case. We therefore assume once again that $\phi_i \in \{0, \infty\}$. With this assumption, we can write the aggregate estimated elasticity among those who face the tax change is

$$\hat{\varepsilon}_1 = \frac{d \log h}{d\,(1-\tau)} = \delta\varepsilon + (1-\delta)\,\zeta\varepsilon \tag{13}$$

Notice that the two terms in this expression are identical to those in equation 11. Intuitively, the two terms play exactly the same role as the individual bunching and firm bunching above. The first term represents the response driven by workers who have no adjustment costs, and thus individually move to their new optimal job. The second term reflects those who face the tax change but cannot respond themselves, and therefore move only to extent that firms change the aggregate job offer distribution. Equation 13 shows then that the estimated elasticity is increasing in the fraction of workers affected. Intuitively, when most workers experience the tax reform, many firms will respond. Even those workers who cannot adjust individually will mostly adjust as a result of the firm responses, and so the estimated elasticity will be large. On the other hand, when very few workers experience the tax reform, very few firms will respond, and so nearly all of the response must come through individuals. The adjustment cost has more bite. Both terms in equation 13 are increasing the underlying elasticity, generating the result in Proposition 3 that firm-driven and individual-driven response to tax changes will covary positively.

*Long-Run Effects of Taxes.* Why should we be interested in identifying the structural elasticity $\varepsilon$ rather than just the observed response $\hat{\varepsilon}$? We now show that the structural elasticity continues to control the long run effects of taxes even with frictions. It is therefore the more relevant parameter for optimal tax policy in the long run.

Consider the same tax reform as that above, which raises the tax rate (in a linear schedule) from $\tau$ to $\tau' > \tau$ for those fraction $\zeta$ facing tax schedule $s = 1$. We showed above that such a tax change in the short run generates response $\hat{\varepsilon} \leq \varepsilon$.

In the long run, however, new generations of workers are born and optimize with respect to the new tax rate $\tau'$ to begin with. Firms shift to reflect the new tax preferences of the population, and fraction $\zeta$ of the jobs in the economy thus shift as well. Any workers who pay to shift do so with respect to the new tax system, and locate at the new optimum. Furthermore, for workers not paying to search, the bounds $\left[\underline{h}_i, \bar{h}_i\right]$ shift; using equations (5) and (6), and an approximation that the bounds are narrow, one can show that the movement in the bounds is also determined by $\varepsilon$:

$$\frac{\partial \ln \underline{h}_i}{\partial \ln (1-\tau)} = \frac{\partial \ln \overline{h}_i}{\partial \ln (1-\tau)} \simeq \varepsilon.$$

Combining these two results, it follows that the observed long run elasticity $\widehat{\varepsilon}^{LR} \simeq \varepsilon$. Intuitively, the choices of the new generation of workers are unaffected by the history of taxes. In contrast, the short-run response of existing workers is attenuated, particularly for small tax changes, because many workers are already close to their new optima and therefore do not pay cost of reoptimizing. Hence, it is important to account for frictions and recover the underlying structural elasticity in order to predict the long-run impacts of taxes on behavior and welfare.[12]

## 3 Data and Institutional Background

We test the model's predictions using data from Denmark because of its substantial variation in tax rates and excellent microdata. The Danish labor market is characterized by what is commonly described as "flexicurity" – low employment protection, large active labor market programs, and high levels of social insurance and transfers. Tax revenues constitute nearly 50% of GDP. Approximately half of the tax revenues fund the provision of public goods and

[12] In more general models of frictions, the long run effect of taxes may differ from and possibly even exceed $\varepsilon$. However, even though the structural elasticity may not fully determine long run responses, it remains a parameter of interest because it affects the long run elasticity.

the other half finance transfer payments.

During the period of study (1994-2001), income was taxed using a three-tier progressive system with a top marginal tax rate of between 63 and 68%. As an example, the 1994 tax schedule is shown in Figure 3a in terms of Danish Kroner (DKr). To facilitate comparisons, note that $1 ≈ DKr 6. The marginal tax rate begins at approximately 45%, which is referred to as the "bottom tax."[13] At incomes of approximately DKr 125,000 and DKr 175,000 , two "middle taxes" are levied on top of the bottom tax. These middle tax brackets lead to reductions in the net-of-tax rate of approximately 9% and 6%. Finally, at incomes above DKr 240,000, one must pay the "top tax" on top of the other taxes, which brings the marginal tax rate to nearly 70%. The net-of-tax wage rate falls by 30% at the point where the top bracket begins. Approximately 25% of individuals pay the top tax during our sample period. The substantial change in marginal tax rates in a central part of the income distribution makes the Danish tax system a particularly interesting case to study for our purposes.

Figure 3b plots the movement in the top bracket cutoff over the sample period in real and nominal terms. Danish tax law stipulates that the movement in the top tax bracket from year $t$ to year $t+1$ is a pre-determined function of wage growth in the economy from year $t-2$ to year $t-1$ (two-year lagged wage growth). This is important for our analysis because it ensures that tax policy is set *before* labor contracts are signed. Over the sample period, inflation was between 1.8% and 2.9% per year. Because of the adjustment rule, the top bracket cutoff declines in real terms in the first three years of the sample and then increases sharply for the remainder of the sample. This variation in the top tax cutoff relative to inflation and contemporaneous wage growth rates proves to be quite useful for identification.

There is some small legislated variation in tax rates across the years in our sample as well as some cross-sectional variation across municipalities and regions. The most notable change in the tax system over the period studied is the consolidation of the two middle taxes into a single tax in 1996. This reform created a single middle kink where the net-of-tax wage fell by 10%. The 1998 revision of the tax code brought other changes. In particular, negative net capital income was excluded from the middle tax base and own contributions to capital

[13] Individuals with very low incomes are exempt from this bottom tax. We restrict attention to individuals with incomes above DKr 50,000 (US $8,300) to avoid dealing with the complications in incentives created by the transfer system at the bottom of the income distribution.

pension plans could no longer be deducted in the top tax base, creating changes in marginal labor income tax rates for many individuals. Finally, during the period from 1994-2001, the middle and top tax bracket cutoffs move both up and down in real terms, creating variation in marginal tax rates for some individuals.

The Danish tax system is in most respects a joint system for couples when calculating the bottom and middle taxes. However, top tax payments are a function of only individual earnings in the vast majority of cases. Capital income enters the tax base for the bottom and middle tax additively, but only enters the top tax base if it is positive.[14] Since mortgage and other interest payments are fully deductible in capital income, most people in Denmark have negative net capital income on their tax return. Thus, the marginal dollar of capital income is effectively untaxed in the top bracket. Most pension contributions are also not taxed. Together, these features of the tax code create an incentive for high earners to shift earnings from labor income to capital income and pensions, responses that we address in our empirical analysis.

There are two tax bases relevant for our empirical analysis: one for the top tax and one for the middle taxes. The top tax base is defined as the sum of earnings, other non-capital income, net positive capital income above a time-varying floor, minus deductions. Earnings include wages and salaries as reported on tax forms plus the value of taxable fringe benefits (such as the use of a company car). Other non-capital income includes self-employment income, transfers (such as unemployment insurance benefits), gifts, and other grants. Deductions consist primarily of individual contributions to pension funds. The middle tax base is defined as the sum of net personal income and net capital income from 1994-1998. In 1999, the definition of the middle tax base changed to include only net personal income and net positive capital income. For consistency with the prior literature, we use the term "taxable income" to refer to the tax base relevant to a particular tax; for instance, when studying bunching around the top tax cutoff, we use "taxable income" to refer to the top tax base.[15]

*Knowledge of the Tax System.* Individuals' perceptions of the tax system may be an important driver of optimization frictions. In Chetty et al. (2010), we analyze what workers

[14] However, the base for the upper middle tax in 1994 and 1995 does not include capital income.

[15] The Danish tax system includes a technical concept of "Taxable Income." We use the phrase "taxable income" here not to refer to that technical concept, but to the relevant base for a given tax.

know about tax system using an internet survey of 3,300 members of FTF-A, a union representing public and financial sector employees. The survey was conducted from March-April 2009. We asked individuals to report their best guess of the top tax and middle tax cutoffs in the current year. These questions were attached to the end of a longer survey on the members perceptions of the union. These survey responses must be interpreted as purely anecdotal evidence because they are not representative of the Danish population, because the survey was administered only to members of FTF-A and because the response rate for our questions is only 11%.

Figure 3c displays the distribution of respondents' best guess at the level of the middle tax (dashed blue) and top tax (solid red) cutoffs among survey respondents. Each point depicts the fraction of responses in a DKr $30,000$ bin centered around the true cutoffs, which are shown by the vertical lines. Knowledge of the top tax cutoff is considerably sharper than the middle tax cutoff. Approximately 41% of the those surveyed know the top tax cutoff to within DKr $15,000$ of the true level; in contrast, only 27% know the middle tax cutoff to within this range. The median absolute error for the top tax cutoff is DKr 21,200, compared with DKr 31,000 for the middle tax cutoff. The same qualitative pattern is exhibited across all education levels and occupations in the sample.

This anecdotal evidence is consistent with our finding that observed elasticities are larger at the top kink than the middle kink, as well as recent evidence that the information and salience affect behavioral responses to income taxation (e.g. Chetty and Saez 2009).

*Data.* We merge several administrative registers provided by Statistics Denmark. The primary data set we use is the tax register, which contains all individuals' tax records along with individual identifiers from 1980 to 2005. This data set contains information on wage income, self-employment income, pensions and pension contributions, capital income and deductions, municipality of residence, filing status, and spouse ID. We then merge the tax returns data with the Danish Integrated Database for Labor Market Research (IDA). The information available in IDA includes education, a firm ID number, occupation, labor market experience, and number of children by age band. Using these data, we construct a tax simulator (code available upon request) that calculates tax liabilities and marginal tax rates for each individual in the sample.[16] The tax simulator predicts actual tax liabilities within 5 DKr for over 97% of wage earners in the sample.

We construct our analysis dataset from a panel of all tax returns for the Danish population between 1994 and 2001. Starting from this population file of approximately 30.1 million observations, we exclude individuals younger than 15 and older than 70, and those who report positive self employment income, pension income, or other unclassified income. These exclusions leave us with a primary analysis sample that consists of 19.4 million observations of wage earners. We also study the population of 1.7 million self-employed individuals as a supplement.

Table 1 presents summary statistics for the population as a whole, our primary estimation sample of wage earners, and the group of self-employed individuals. The mean taxable income for an individual in the sample in real 2000 kroner is DKr 185,749, or approximately $32,000. The median taxable income is DKr 176,217. The much smaller difference between mean and median incomes in Denmark than in the U.S. reflects the more compressed income distribution in Denmark.

## 4 Empirical Analysis

We study the impact of frictions on the relationship between taxes and labor supply by testing the three predictions derived in Section 2. Section 4.1 provides basic evidence of behavioral responses to the tax system, showing that larger variation in net-of-tax rates generates disproportionately larger responses. This section also presents some auxiliary evidence which suggests that the observed taxable income responses arise from "real" changes in behavior rather than reporting effects. Section 4.2 shows that much of the bunching at kinks is driven by coordinated setting of salaries by firms, demonstrating the importance of endogenous hours constraints. Section 4.3 examines the heterogeneity in individual bunching across occupational

[16] The source variable for personal income from the tax data is *perindkp*. The source variable for captial income is *kapindkp*, and the source for deductions to taxable income is *lignfrdp*. The source variable for pension contributions is *kappens*. Individual contributions to pensions are considered deferred income for tax purposes, and thus are included as deductions in personal income. However, starting in 1999, deductions for own contributions to capital pension schemes are added back again when calculating the top tax base. Using a person's filing status, *samskat,* and the social security number of spouses, *henv*, we account for rules regarding the transfer of exemptions, negative capital income, etc. between spouses.

and demographic groups and links this heterogeneity to variation in search costs.

### 4.1 Prediction 1: Size of Tax Changes

There are two sources of tax variation in the data: legislated changes in tax rates across years (tax reforms) and variation in tax rates across tax brackets. We first examine the cross-sectional variation in tax rates across brackets, which is much larger than the variation created by reforms. If one allows an arbitrary smooth wage distribution, the only non-parametric source of identification in the cross-section is the degree of bunching at kink points (Saez 2002). Hence, when exploiting cross-sectional variation in tax rates, we focus exclusively on estimating the amount of bunching at kinks in the tax schedule. Our focus on bunching at kinks contrasts with the approach taken in the non-linear budget set literature, where the variation in tax rates at kinks is viewed as being problematic for identification. This is because there is very little bunching at the small kinks examined in prior work, leading to negative estimates of compensated wage elasticities that are inconsistent with utility maximization (MaCurdy, Greene, and Paarsche 1990). This forces researchers to remove or ignore the variation at the kinks, e.g. by "smoothing" the budget set. As our model provides a microfounded explanation for why some kinks generate much more bunching than others, we are able to exploit the kinks to our advantage rather than cast them aside as problematic variation.

We begin our analysis by examining bunching at the top bracket cutoff, which is the largest kink in the tax schedule.[17] We then examine bunching at smaller kinks, showing that this tax variation produces much smaller behavioral responses. Finally, we investigate the effects of tax reforms, which generate smaller variation in tax rates and produce estimates consistent with those obtained from the smaller kinks. We use the tax reforms to distinguish the effects of adjustment costs on observed responses to kinks of different sizes from preferences that lead to non-constant elasticities.

[17] Strictly speaking, the Danish tax system includes a larger kink between DKr $20,000$ and DKr $30,000$ when individuals exhaust the personal exemption and begin paying tax. In practice, the welfare system greatly complicates the incentives to work in this income range. We therefore do not analyze behavior around this kink.

### 4.1.1 Bunching at the Top Bracket Cutoff

When testing the first two predictions, we drop individuals who report positive self-employment income and focus exclusively on wage earners.[18] Adjustment costs and institutional frictions are much less likely to hamper adjustment of self-employment income than wage earnings. Correspondingly, we show in Section 4.3 that the effects of taxes on self-employment income are an order-of-magnitude larger than on wage earnings.

Figure 4 plots the empirical distribution of taxable income for all wage earners from 1994-2001. To construct this figure, we first calculate the difference between the actual taxable income and the taxable income that would be needed to reach the cutoff for the top tax bracket for each observation in the sample. We then group individuals into DKr 1,000 bins on this recentered taxable income variable. Finally, we plot frequencies of the observations in each of the bins to obtain a non-parametric depiction of the empirical distribution of income around the top bracket cutoff, which is demarcated by the vertical line at zero. An individual's marginal wage falls on average by 30% when his income crosses this vertical line.

Figure 4 shows that there is a clear spike in the density around the top bracket cutoff in the otherwise smooth and monotonically declining distribution. This spike indicates that the sharp reduction in marginal wage rates in the top tax bracket induces many individuals to reduce their taxable incomes. Following Saez (2009), we infer the elasticity of taxable income with respect to the net-of-tax rate from the amount of excess mass around the kink. To do so, we predict a counterfactual density – what the distribution would look like if there were no jump in tax rates – using the following procedure. First, we fit a polynomial to the counts plotted in Figure 4, *excluding* the data that is near the kink. Formally, we estimate a regression of the following form:

$$C_j = \sum_{i=0}^{q} \beta_i (y_j^i) + \sum_{l=-r}^{r} \gamma_l \mathbf{1}\{y_j = l\} + \varepsilon_j \tag{14}$$

where $C_j$ are the number of individuals in income bin $j$, $y_j$ is income relative to the kink in 1,000 Kroner intervals (i.e., $y_j = \{-50, -49, .., 50\}$), $q$ is the order of the polynomial to be fit,

[18] We check that the endogenous sample selection induced by dropping those who report self-employment income does not spuriously generate bunching by verifying that there is significant bunching in the wage earnings distribution in the full sample.

and $r$ denotes the width of the excluded region around the kink (measured in DKr 1,000 ). Although the choice of the degree of the polynomial is arbitrary, numerical simulations indicate that a seventh-degree or higher polynomial ($q = 7$) and exclusion of $7,000DKr$ around the kink ($r = 7$) accurately capture the degree of bunching at the kink for distributions similar to that in Figure 4. Moreover, for every calculation reported below, we have verified that the results are not sensitive to changes in $q$ and $r$, presumably because of the sharpness of the bunching patterns we observe in the data. We define the counterfactual density as the predicted values of this regression omitting the contribution of the dummies around the kink, i.e., $\widehat{C}_j = \sum_{i=0}^{q} \widehat{\beta}_i y_j^i$. The counterfactual density predicted using this procedure is shown by the solid line in Figure 4. The excess number of individuals who locate at the kink relative to the counterfactual density is $b_n = \sum_{l=-r}^{r} \widehat{\gamma}_l$, the shaded area in Figure 4. The excess fraction around the kink is

$$b_f = \frac{b_n}{[\sum_{l=-r}^{r} \widehat{C}_l] - b_n}$$

In Figure 4, $b_f = 5.97\%$ implying that there is 5.97% more mass around the kink in the actual empirical distribution relative to the counterfactual distribution predicted if there were no higher bracket in the tax system.

This method of calculating the counterfactual density overestimates $b_f$ because it does not account for the fact that the additional mass at the kink comes from points to the right of the kink. The individuals at the kink reduce their earnings because of the higher tax rate in the top bracket; absent the kink, they would locate to the right of the vertical line at zero and therefore increase the counterfactual density. One approach to addressing this error is to use a parametric model of utility maximization to reallocate the individuals to the right of the kink. An alternative is to gauge the degree of error from the simpler non-parametric approach in equation (14). In Appendix A, we provide a calibration showing that our procedure overestimates $b_f$ by approximately 5-10%. Because the sharp, visually evident variation in $b_f$ that we document below is unlikely to be affected by this degree of bias, we opt to use the simpler and more transparent non-parametric approach of calculating $b_f$ above.

The standard errors associated with the regression estimates of $\beta$ and $\gamma$ in equation (14) yield standard errors for $b_n$ and $b_f$. Since we observe the population distribution of income, these standard errors reflect error due to misspecification of the polynomial for the counterfactual income distribution rather than sampling error. In Figure 4, the standard error associated with $b_f$ is $b_{se} = 0.38\%$. Hence, the null hypothesis that there is no excess mass at the kink relative to the counterfactual distribution is rejected with a t-statistic of $5.97/0.38 = 15.7$ (implying $p < 1 \times 10^{-9}$). The reason that the standard error $b_{se}$ is so small is that the polynomial approximates the smooth empirical income distribution extremely well, as indicated by the tightness of its fit away from the kink. Hence, there is little specification error in the counterfactual, and one can conclude with confidence that there is a spike in the income distribution at the top tax bracket cutoff.

Figure 4 masks a substantial amount of heterogeneity across groups in responsiveness to the jump in tax rates. Figure 5a shows that the excess mass at the kink is much larger for married women ($b_f = 14.1\%$), consistent with the general perception that married women's labor supply is the most elastic to incentives. Conversely, some groups of the population exhibit much less responsiveness to the tax system. For example, $b_f = 1.83\%$ for single men, as shown in Figure 5b. Married women who have more than the median level of labor market experience (19 years) are especially elastic: Figure 5c shows that $b_f = 18.7\%$ for this group. The degree of bunching is even larger if we focus on high-experience women who are white collar professionals: $b_f = 41.5\%$ as shown in Figure 5d. We return to explore these and other dimensions of heterogeneity in Section 4.3 below. In the remainder of this section, we focus on married women, as this is the group for whom behavioral responses to the tax system are clearest; however, all of the qualitative points made below hold in the full sample of wage earners.

The identification assumption underlying causal inference about the effect of taxes on earnings in the preceding analysis is that there would be no spike in the income distribution if there were no jump in tax rates at the location of the top bracket cutoff. This identification assumption can be relaxed by exploiting the movement in the top bracket cutoff across years to test whether the excess mass moves with the cutoff. Figure 6 displays the distribution of taxable income in each year from 1994-2001 for married women. There is sharp bunching at the top bracket cutoff in every year. The excess mass follows the top bracket cutoff very closely, which is informative particularly because the bracket cutoff declines in real terms from 1994-1997 and then increases sharply until 2001, as shown in Figure 3b.

In Figure 7, we investigate more directly whether the location of the bunch follows the top tax cutoff, inflation, or average wage growth. Figure 7a plots the distribution of taxable income in 1997 (as in Figure 6d). Given that the bunch locates at the top tax cutoff in 1994, we mark with vertical lines three possibilities for the location of the bunch in 1997: the 1997 top tax cutoff (dashed line), the 1994 cutoff adjusted for inflation (left-most solid line), and the 1994 cutoff adjusted for average wage growth (right-most solid line). The excess mass at the 1994 kink clearly moves to the 1997 kink rather than following inflation or wage growth. Figure 7b replicates Figure 7a for the year 2001; the solid vertical lines show the 1997 top tax cutoff adjusted for inflation and wage growth, respectively. Again, the excess mass clearly follows the top tax cutoff rather than wage growth or inflation. Hence, earnings around the top tax bracket depart from prevailing inflation patterns and instead are aligned with changes in the tax system.

#### 4.1.2 Shifting vs. Real Responses

Individuals can reach a taxable income near the top bracket cutoff through two margins: changes in labor supply (e.g. hours worked) or "income shifting" responses such as changes from taxed to untaxed or tax-favored forms of compensation. It is useful to distinguish between these two channels because shifting of income can potentially have very different normative implications than "real" changes in labor supply (Slemrod 2002, Chetty 2009b, Saez et al. 2009). To identify income shifting effects empirically, it is useful to conceptualize the various channels through which taxes can affect reported taxable income and total compensation. An individual's total compensation is:

$$\begin{aligned} \text{total comp.} \quad = \quad & (1-\tau)(\text{wage} \times \text{hours} - \text{evasion} - \text{pension contribs.} + \text{other taxable comp.}) \\ & + \text{pension contribs.} + \text{evasion} + \text{unobserved nontaxable comp.} \end{aligned}$$

Total compensation is the sum of salaries and other taxable compensation, such as in kind payments, and untaxed compensation (contributions to untaxed pension accounts and other unobservable compensation such as workplace amenities). Total compensation is also affected by tax evasion, which reduces taxed compensation and increases untaxed compensation, leading to a net gain of $\tau$ times the amount evaded.

This expression shows that aside from a change in hours (labor supply), there are two income shifting channels through which an increase in $\tau$ may induce individuals to change their reported taxable income. First, individuals may simply under-report earnings via tax evasion. Kleven et al. (2009) conduct an audit of wage earnings in Denmark and conclude that there is virtually no tax evasion in the line on the tax form from which our wage earnings variable is constructed. Kleven et al. argue that tax evasion in wage earnings is very uncommon because wage earnings are double-reported to the government by both firms and individuals. To ensure that the bunching we see is not the result of evasion or underreporting, we examine the distribution of pure wage earnings, *excluding* all other forms of compensation. Figure 8a plots the empirical distribution of wage earnings for married women, and shows that there is clear bunching even in this narrowest measure of compensation. This evidence shows very clearly that individuals are changing their wage earnings in order to stay near the top bracket cutoff and not some other component of compensation. Indeed, there is more bunching in wage earnings ($b_f^{wage} = 18.4\%$) than in taxable income (where $b_f = 14.1\%$), a surprising result that we explain in Section 4.2. We therefore conclude that the changes in taxable income that we see around the top bracket cutoff are not driven by tax evasion.

The second and more important income shifting channel is legal tax avoidance: individuals may shift compensation toward nontaxable compensation. The simplest method of avoiding the top income tax is to increase the amount contributed to non-taxable pension accounts (similar to 401(k)s). We investigate the extent of such shifting by adding pension contributions to taxable income.[19] Figure 8b plots the distribution of this broader measure of compensation relative to the top tax bracket cutoff that applies to individuals with *zero* pension contributions. There is still clear evidence of bunching at the top kink, rejecting the hypothesis that all of the bunching observed in taxable income is driven by shifts to pension contributions. The excess mass around the top bracket is now smaller than in Figure 5a: $b_f = 8.72\%$ for pensions plus taxable income, compared with $b_f = 14.1\%$ for pure taxable income. This is not surprising because the vertical line at zero no longer represents the point at which tax rates jump for individuals who have non-zero pension contributions – for such individuals, there is no reason to locate at this point even if they are only changing labor supply. To quantify how much

[19] In Denmark, there are two different types of pension schemes to which workers can make tax-free contributions. One scheme is run through the firm, as is a 401(k); the other is a private tax-deductible pension, as is an IRA. We account for both forms of pension contribution in this analysis.

of the shrinkage in the excess mass is due to this mechanical attenuation effect, we estimate the fraction of individuals who are pulled out of the band around the top bracket (whose width is $2r = 14,000DKr$) by their pension contributions. Using data from individuals with incomes between $20000DKr$ and $50000DKr$ — a range which places them well away from the top tax bracket cutoff – we calculate that 64.5% of individuals make pension contributions above DKr 7,000. Correcting for the mechanical attenuation effect therefore yields $b_f^{TI+pens} = 8.72\%/(1-64.5\%) = 13.5\%$. The remaining difference in $b_f^{TI+pens}$ and $b_f$ for the pure taxable income measure ($14.1\% - 13.5\% = 0.6\%$) reflects the amount of bunching driven by pension shifting. which of course we should expect, to some extent, if workers respond optimally to the tax system. We conclude that pension shifting accounts for a very small fraction of the observed response in taxable income.

Another margin through which individuals can shift income to avoid the top tax is through capital income. Capital income (or losses) do not enter the tax base in the top bracket for more than 3% of the population. Hence, most individuals can reduce their tax liabilities by shifting compensation from wage income into capital income when they cross into the top tax bracket. The distribution of taxable income plus pensions plus capital income (CI) also exhibits excess mass around the top kink (not reported). Applying the correction for mechanical attenuation due to non-zero capital income described above, we find that $b_f^{TI+pens+CI} = 11.4\%$. Again, shifting into capital income accounts for a small fraction of the changes in taxable income that we observe.

Our analysis shows that bunching is not driven by any of the *observable* methods of income shifting, but we cannot rule out the possibility that individuals shift their compensation to unobservable nontaxable compensation to avoid paying the top income tax. For example, we cannot detect substitution of compensation from wage earnings into free lunches or nicer offices when individuals cross into the top tax bracket. The only means of definitively ruling out such responses is to examine direct measures of "labor supply", such as hours worked. Unfortunately, such measures are unavailable in our data, and are in any case typically quite poorly measured, as reported hours likely do not capture most of the variation in actual work effort (e.g., length of breaks, vacations).

Although we cannot conclusively distinguish between shifting and real labor supply changes using the available data, in view of the evidence above, we believe that pure shifting responses

– of the type that have little or no efficiency cost – account for a small fraction of the observed bunching in taxable and income, for three reasons. First, the most straightforward and accessible means of legally shifting income would be changing pension contributions; it would be surprising if individuals resort to more exotic means of shifting income when they do not use this channel. Second, few individuals at the 75th percentile of the income distribution have the ability to shift income into other forms of compensation or across time. Existing evidence shows that income shifting responses are quite large only among very high incomes, who can restructure their salary packages in a sophisticated manner to avoid taxes (Slemrod 2002). Third, it is clear that if shifting is occurring, it is not through costless relabeling of taxable income as another untaxable form of cash compensation. If shifting occurs through costly channels such as compensation in free lunches and office perks instead of cash compensation, then the marginal efficiency cost of such behaviors equals the marginal efficiency cost of changes in labor supply itself in an equilibrium where agents optimize (Feldstein 1999). In this sense, the welfare consequences of the changes in taxable income that we document are likely to be similar to the welfare consequences of changes in labor supply.

Finally, note that the income shifting vs. labor supply distinction is ancillary to the main goal of the present paper, which is to investigate how adjustment costs and institutional rigidities affect behavioral responses to taxation. Whether the responses we document occur through changes in hours or other costly income shifting channels, the evidence below shows that it is important to take adjustment frictions into account in order to understand the impacts of tax policies on the economy.

### 4.1.3 Bunching at Smaller Kinks

To test prediction 1, we measure the amount of bunching at other kinks in the Danish tax system that create smaller changes in the net-of-tax rate. Note that even without adjustment costs, one would predict less bunching at smaller kinks, simply because the change in incentives is smaller. We therefore compare the amount of excess mass at these smaller kinks with the amount of excess mass that would be generated if the elasticity were the same as that implied by the excess mass at the large top tax kink examined above. To do so, we take the observed distribution of earnings around these smaller kinks, after removing the observed kink itself, and calculate the implicit distribution of wages given a linear tax and using the elasticity consistent with bunching at the top kink.[20] We then simulate optimal labor supply for individuals around the smaller kink and plot the counterfactual distribution against the actual distribution. This allows us to assess visually whether behavior departs from a model with no adjustment costs and constant-elasticity utility as in Section 2.

We examine the taxable income distribution around three "middle tax" brackets, which generate variation in tax rates of different sizes across the years in our sample. Between 1994 and 1995, the tax system also included an "upper middle tax" of 5% in 1994 and 3% in 1995 that was levied on incomes above DKr 173100 and 174300, respectively. Figure 9a plots the empirical and simulated income distributions around this tax bracket, where net-of-tax wages fall by 6% at the vertical line. Again, there is essentially no bunching at this kink in the actual distribution ($b_f = 0.98\%$, $b_{se} = 0.90\%$). In contrast, the simulated distribution based on behavior at the top kink implies that we should see $b_f = 4.8\%$ absent adjustment costs. Between 1994 and 1995 the tax system also included a "lower middle tax". The tax rate for this was 4.5% in 1994 and 5% in 1995 and was levied on incomes above DKr 130000 and 130900 respectively. Figure 9b plots the empirical income distribution around this tax bracket in blue circles, using the same methodology as above. The estimated $b_f = -0.44\%$ ($b_{se} = 0.47\%$), and the upper bound of the 95% confidence interval is 0.50%. There is virtually no response to the 6% reduction in the net-of-tax wage at the vertical line. For comparison, the distribution plotted in squares shows the simulated distribution around this kink if the elasticity were the same as what is implied by behavioral responses around the top kink. It is clear that the actual behavioral response to the lower middle kink is much smaller than what one would predict based on behavior around the top kink in a model without adjustment costs.

Finally, between 1996 and 2001, the lower and upper middle taxes were consolidated into a single tax of 5% in 1996 and 6% thereafter, that was levied on incomes above DKr 134500, 135300, 139000 for the years 1996-1998 and increased markedly to DKr 151000, 164300, and 177900 between 1999 and 2001. Figure 9c plots the income distributions around this tax bracket, where net-of-tax wages fall by 10%. There is evidence of modest bunching in the actual distribution around this slightly larger kink, with $b_f = 2.24\%$ and $b_{se} = 0.46\%$. However, the observed level of bunching still falls well below the simulated amount with constant

[20] We use the isoelastic quasilinear utility function, as in the theory section above, for this exercise.

elasticities.

#### 4.1.4 Estimating Elasticities from Small Tax Reforms

We now turn to the second source of variation in tax rates: changes in marginal rates by legislated reforms. As described in Section 3, there were a number of small tax reforms in Denmark during our sample at both the national and regional levels. These tax reforms create changes in net-of-tax rates of between -5% and +7%. These changes have different effects across the income distribution because of the progressivity of the tax system. This differential variation across income groups motivates the use of a difference-in-difference research design that is standard in the taxable income literature (Saez et al. 2009). We estimate an observed elasticity using instrumental-variable specifications of the same form as Gruber and Saez (2002) and Kleven and Schultz (2009). Let $\Delta \log y_{it}$ denote the log change in wage earnings from period $t-2$ to period $t$ and $\Delta \log(1-MTR_{it})$ the log change in net-of-tax rates over the same period. We estimate the following model using two-stage least squares:

$$\Delta \log y_{it} = \alpha + \beta \Delta \log(1 - MTR_{it}) + f(y_{it}) + \gamma X_{it} + \varepsilon_{it}, \tag{15}$$

instrumenting for $\Delta \log(1 - MTR_{it})$ with $\Delta \log(1 - MTR_{it}^{\text{sim}})$, the simulated change in net-of-tax rates holding the individual's income and other characteristics fixed at year $t$ levels. The function $f(y_{it})$ is a 10-piece linear spline in base year wage earnings and the vector $X_{it}$ is a set of base year controls that we vary across specifications. The first-stage regression of $\Delta \log(1-MTR_{it})$ on $\Delta \log(1-MTR_{it}^{\text{sim}})$ has coefficients of approximately 0.9 with t-statistics exceeding 1000; as a result, the two-stage-least-squares estimates are very close to estimates from reduced-form regressions of $\Delta \log y_{it}$ on $\Delta \log(1 - MTR_{it}^{\text{sim}})$.

We estimate variants of (15) for the subsample of female wage earners because the likelihood of detecting responses to small reforms is larger for more elastic groups. Figure 10 illustrates the variation that identifies these regressions using a scatter plot of residuals. To construct this figure, we first regress $\Delta \log(1 - MTR_{it}^{\text{sim}})$ on the 10-piece wage earnings spline and the following controls: a 10-piece spline in total personal income and year, age, occupation, and region fixed effects. We then calculate residuals from this regression. Similarly, we regress $\Delta \log y_{it}$ on the same set of covariates and compute residuals. Finally, we bin the residual changes in simulated net-of-tax rates into 0.1% intervals. The figure plots the mean residual change in wage earnings in each bin. As a visual guide, the figure also shows the estimated relationship between the residuals (red solid line), which can be interpreted as a rough estimate of the observed elasticity. The grey shaded region shows the 95% confidence interval for this estimate of the reduced-form relationship between changes in simulated net-of-tax rates and changes in earnings. The figure shows that changes in net-of-tax wages of between -5% and +6% have no effect on observed wage earnings. Even individuals who experience changes in net-of-tax wages of more than 5% in have residual growth rates confined between $-1.5\%$ and +1%. The slope of the fitted line (red solid line) implies an observed elasticity of $\widehat{\varepsilon} = 0.00$. The 95% confidence interval spans a width of 0.2%, implying an upper bound on the observed elasticity of $\widehat{\varepsilon} \simeq 0.02$.

Building on the identification strategy illustrated in Figure 10, we report TSLS estimates from several variants of (15) in Table 2 for female wage earners. In column 1, we estimate (15) on the sample of all female wage earners with the full set of controls used in the regressions used to construct Figure 11. Consistent with the graphical evidence, the estimated elasticity is $\widehat{\varepsilon}$ is not significantly different from 0, and the upper bound of the 95% CI is 0.002. In column 2, we explore the robustness of the results to changes in the covariate set by including only the wage earnings spline and age and year fixed effects as controls. The estimated elasticity remains very close to zero, showing that the estimates are robust to the set of covariates used to predict counterfactual income growth. In columns 3-5, we focus on subsamples where we observe larger elasticities using the large variation in tax rates at the top kink and estimate specifications with the full set of controls used in column 1. Column 3 considers married females (whose propensity to bunch at the top kink is illustrated in Figure 5a), column 4 considers married female white collar professionals with high experience (Figure 5d), and column 5 teachers (Figure 14a below). In each of these groups, the observed elasticities in response to small tax changes remains very small and not significantly different from 0. Hence, estimating taxable income elasticities using small tax reforms uncovers reveals no behavioral response to taxes in Denmark, consistent with the prior literature.

#### 4.1.5 The Effect of the Size of Tax Changes on Observed Elasticities

Figure 11 compiles the evidence on observed elasticities and the size of tax changes. It plots the observed elasticities for married women ($\hat{\varepsilon}$) vs. the change in the net-of-tax wage ($\frac{\Delta t}{1-t}$) at the three kinks described above, and also includes the elasticity estimate using small tax reforms from Column 3 of Table 2. To convert the observed excess masses to elasticity estimates, we apply the formula in equation (1) from Saez (2009). Recall that a model without adjustment costs and constant elasticity utility as in Section 2.2 predicts that $\hat{\varepsilon}$ does not vary with $\frac{\Delta t}{1-t}$. To test this hypothesis, we fit a linear regression through the five elasticity estimates. The hypothesis that the slope of the regression line equals zero is rejected with $p < 0.01$.

Even though there is substantially more bunching at the largest kink, the implied elasticities even at this kink remain small. Applying the formula in equation (1) from Saez (2009) for married women, $\widehat{\varepsilon} \simeq 0.03$ at the 30% tax kink. The structural (long run) elasticity is likely to be much higher than this estimate because (as Figure 2a shows) even a 30% change in wages might be insufficient to induce all individuals to reoptimize their labor supply choices and locate at the kink. Chetty (2009a) calculates that the utility gains from reoptimizing relative to a 30% tax change are only 2% of consumption if the structural elasticity $\varepsilon = 0.5$. Individuals whose adjustment or search costs exceed 2% of consumption will therefore not respond to even the largest tax changes in our data despite having a true long-run elasticity of $\varepsilon = 0.5$. In Section 5 below, we calibrate the model discussed above and show that, given an upper bound of 20% on the size of the adjustment cost, the data place a lower bound of 0.28 on the structural elasticity.

#### 4.1.6 Distinguishing Adjustment Costs from Non-Constant Elasticities

The conclusion that the upward-sloping relationship in Figure 11 is driven by adjustment costs is predicated on the assumption that the structural elasticity $\varepsilon(\tau, z)$ remains constant as the tax rate $\tau$ and agent's income $z$ changes. If $\varepsilon(\tau, z)$ varies with $\tau$ or $z$, the differential responses we observe at the various kinks could be generated by variations in $\varepsilon$. We now turn to distinguishing this alternative explanation of our findings from the adjustment cost mechanism.

There are two channels through which $\varepsilon$ could vary with $\tau$ or $z$. First, elasticities may be heterogeneous across individuals. For instance, higher income individuals might have more elastic preferences. In this case, the larger amount of bunching at the top kink than the middle kinks could be driven by preference heterogeneity that is correlated with the tax rates. Second, a given individual's utility function may be such that the local elasticity $\varepsilon(\tau, z)$ is an increasing function of $\tau$ or $z$. This would again lead to more bunching at the top kink than the middle kinks for reasons unrelated to adjustment costs.[21]

As shown in Section 2, the effects of adjustment costs can be non-parametrically distinguished from variation in $\varepsilon(\tau, z)$ by identifying the sum of the effects of several small tax changes that span a single larger change. In our application, the large change is the increase in the marginal tax rate from approximately 58% to 70% in earlier years and from 51% to 62% in later years at the top kink. The non-parametric test in Section 2.6 would require smaller tax changes that fully span the range from 58% to 70%.

Although we do not have adequate tax variation to implement this non-parametric test, we can approximate it by examining small reductions in the top tax rate. Intuitively, if the local elasticity $\varepsilon(\tau, z)$ around the top kink is larger than $\varepsilon(\tau, z)$ around the middle kinks, then a reduction in the top tax rate from 70% to 65% should generate large behavioral responses. If the difference in tax responses we identified above is instead caused by adjustment frictions, the same tax cut should not generate a significant response.[22] Regressions using small variations in the top marginal tax rate analogous to those in Section 4.1.4 remain very close to zero (not reported), as in the full sample. Hence, there is no evidence that the structural elasticity is itself higher for individuals close to the top bracket cutoff.

As a supplementary test of whether preference heterogeneity drives the differential bunching at the middle and top kinks, we focus on a subset of individuals whose incomes place them within DKr $25,000$ of the top kink in year $t$ but within DKr $25,000$ of one of the middle kinks

[21] A further confound is the possibility that the amount of uncertainty falls with the income level. We think this is implausible because the importance of variable forms of compensation (bonuses etc.) actually increase with income levels.

[22] This test is an approximation of the test in Section 2.6 because it relies on the parametric assumption that the average elasticity $\varepsilon(\tau, z)$ for $\tau \in (65\%, 70\%)$ equals the average elasticity for $\tau \in (58\%, 70\%)$. In principle, it is possible that the elasticity is small below 58%, large between 58% and 65%, and small between 65% and 70%. Such preferences would generate large responses at the top kink, small responses at the middle kink, and small responses to a tax cut from 70% to 65% even in a model without adjustment costs. We believe that such sharp fluctuations in local elasticities are unlikely, and that adjustment frictions are a more plausible explanation of our findings.

in year $t+2$. By studying these "switchers," we can effectively remove individual fixed effects when comparing responses to the middle and top kinks. Figures 12a and 12b display the distribution of taxable income for the group of switchers at each kink. When at the top kink, these individuals exhibit substantial bunching ($b_f = 8.94\%$). However, just two years later, the same individuals show no excess propensity to bunch at the middle kink ($b_f = 1.37\%$) despite having earnings near that kink. The opposite pattern is observed for those moving from the middle to the higher kinks between years.

We conclude that variation in $\varepsilon(\tau, z)$ does not explain the differences in bunching at the middle and top tax kinks. Irrespective of initial tax rates and incomes, small tax changes induce little or no behavioral response, while larger tax changes induce much larger responses. The results are consistent with the hypothesis that adjustment costs attenuate observed labor supply elasticities. In the next two sections, we provide more direct evidence on the mechanisms of adjustment by showing how institutional constraints and variation in search costs affect behavioral responses to kinks in the tax code.

### 4.2 Prediction 2: Firm Bunching and Scope of Tax Changes

In this section, we show that tax variation that affects a larger group of workers has a greater impact on the distribution of wage-hours packages offered by employers. Specifically, we study whether firms set salaries to enable employees to locate close to the top bracket cutoff, which is the optimal earnings level for many workers.

We obtain variation in the size of the group of workers affected by the top tax by exploiting cross-sectional heterogeneity in the size of deductions or unearned income across individuals, which effectively varies the location of the top tax bracket. Recall that taxable income is the sum of wage earnings (paid by the employer) and non-wage taxable compensation minus deductions, as described in Section 3. Non-wage taxable compensation includes pensions received, alimony received, stipends, payments received for being member of a board, and unemployment benefits and severance payments. Pension contributions made by individuals are deducted from taxable income. Because of the variation in non-wage income and deductions, the level of wage earnings required to reach the top bracket varies across individuals. Figure 13 plots the distribution of individuals' "adjustments" to taxable income (deductions minus non-wage income) in 1995. The most common level of adjustments is DKr 0; more than 58%

of individuals have adjustments less than DKr 5000. In this setting, the second prediction of our model is that individuals with small adjustments – who have common tax preferences – should be more likely to bunch at the top tax kink than those with large adjustments.

To characterize employer responses to workers' tax preferences, we focus on wage earnings distributions at the occupation level because most workers' wages are set through collective bargains in Denmark. We first present a case study of salary setting in one of the largest occupations – elementary school teachers. We then show that the lessons illustrated by school teachers generalize to the population as a whole.

### 4.2.1 Case Study: School Teachers

There are approximately 75,000 elementary school teachers in Denmark, constituting approximately 3% of wage earners. Teachers belong to a union that bargains with the state on salaries, work hours, and other job characteristics at a national level. This large-scale collective bargaining facilitates aggregation of worker preferences, providing a setting where the impacts of coordination on responses to tax incentives are especially clear.

Figure 14a plots the taxable income distribution around the top tax bracket for school teachers, as in Figure 4. There is very sharp bunching around the top kink: $b_f = 36.8\%$. How do so many individuals find jobs that place them at the top income tax cutoff? Figure 14b plots the distribution of wage earnings (salaries) for teachers. The distribution of salaries is set so that teachers with zero non-wage income have taxable income exactly at the top tax bracket. Figure 15 shows how teachers' salaries change over time, plotting salary distributions in 1995, 1998, and 2001. The distribution of salaries moves in close correspondence with the top tax bracket rather than the rate of inflation or the average rate of earnings growth in the economy.[23] The small increase in the top tax bracket of DKr $14,600$ between 1995 and 1998 is associated with a small increase in the modal teacher's salary. In contrast, the larger increase of DKr $26,800$ from 1998 to 2001 is associated with equivalently rapid growth in the modal teacher's salary.[24]

[23] One may be concerned that the top tax bracket is set just above the modal earnings levels chosen by unions in each year. This reverse causality explanation is ruled out by the fact that the movement of the top tax bracket is pre-determined by lagged wage growth rates, as discussed above.

[24] The smaller peak visible approximately DKr 15,000 above the kink in each of these figures is driven by teachers in the city of Copenhagen, who are given a cost-of-living adjustment of DKr 15,000 over the base

The intuition for the link we observe between earnings and taxes is straightforward. The rate of return to negotiating for higher wages falls discontinuously for the vast majority of teachers at the top tax bracket cutoff. It is therefore sensible that teachers start bargaining on other dimensions, such as fewer classes, rather than continue to push for wage increases beyond this point.

The salary distribution for teachers matches the tax-induced preferences of the large mass of individuals who have zero non-wage income. Such a salary distribution could arise through two channels. The first is "individual bunching" – each individual searches for a job that suits his tastes. Since many individuals' ideal earnings level coincides with the top tax cutoff, the resulting equilibrium will feature a salary distribution with a mass point at the tax kink. An alternative mechanism is "firm bunching" – firms cater to the tax-induced preferences that are most common in the population and structure contracts (hours and salaries) to locate many individuals at the top kink.

To distinguish between these two mechanisms, we examine teachers whose tax incentives diverge from those of the majority because of non-wage income. Figure 16 plots the distribution of wage earnings around the *statutory* top tax cutoff (that is, the level of taxable income at which the top tax begins) for teachers whose deductions exceeds DKr $20,000$. Note that these individuals do not begin to pay the top tax on wage earnings until at least DKr $20,000$ beyond the level of the top tax cutoff, so there is no change in marginal net-of-tax wages at the vertical line at zero in this group of individuals. Yet the wage earnings distribution for these workers is extremely similar to the distribution for the population as a whole, and exhibits sharp bunching at top tax cutoff. This result shows that firm bunching is prevalent. Intuitively, schools are forced to offer a limited number of wage-hours packages in order to coordinate class schedules. Because of such technological constraints, firms write teachers' contracts to cater to the tax incentives reflected in the population. The tax-distorted preferences of the majority of individuals spill over onto the minority of teachers whose tax incentives differ.

Because of changes in the supply of jobs, teachers with common tax preferences – i.e., those with little non-wage income or deductions – are more likely to respond to the top tax than

teacher's salary. The setting of salaries to place teachers outside Copenhagen – who account for 15% of all teachers – at the top kink supports the view that institutional constraints are endogenously set based on the preferences of the largest groups in the population.

other teachers. This is illustrated in Figure 17, which plots the distribution of taxable income for teachers with non-wage income or deductions less than $5,000$ – whose personal tax cutoff is within DKr $5,000$ of 61% of the population – and teachers with deductions greater than $20,000$. Teachers who have deductions greater than DKr 20,000 exhibit clear excess mass at the kink. Since these individuals cannot rely on changes in the supply of jobs to place them at the kink, this reflects individual bunching. But the excess mass at the kink is much larger for teachers with low non-wage income, confirming the prediction that taxes which affect a larger mass of individuals generate larger elasticities.

As a final test of the effects of firm-induced responses to taxes, we study the dynamics of earnings as the top tax bracket moves over time. To study earnings dynamics, we define an indicator for whether an individual's change in earnings from year $t$ to year $t+2$ is within DKr 7,000 (the width of our bunching window) of the change in the top tax bracket cutoff from year $t$ to year $t+2$. This indicator measures whether an individual's earnings tracks the movement in the kink over time. Figure 18 plots the fraction of individuals who track the movement in the kink by taxable income in the base year for teachers with non-wage income less than DKr 10,000 (firm bunchers) and more than DKr 10,000 (individual bunchers). For firm bunchers, the propensity to track the movement in the kink is maximized for individuals near the kink. Conditional on being near the kink in year $t$, teachers with common tax preferences have a probability of persisting near the kink in year $t+2$ of more than 60%, compared with 20% for teachers with uncommon tax preferences who rely on job search to locate at the kink in year $t$. Intuitively, an increase in the top tax bracket creates a local change in tax incentives for the large pool of teachers who have little non-wage income. As a result of the tax reduction, more teachers are willing to work more in exchange for a higher salary. Firms change the contracts that they offer to accommodate the changes in the average teacher's preferences. In contrast, teachers who locate at the top kink individually must find a new, higher paying job that places them at the new kink. This search cost reduces the responsiveness of individual bunchers to changes in local tax incentives. From Figure 17, we know that there is substantial individual bunching at the kink; but Figure 18 shows these few of these individual bunch for many periods. The central role of changes in the supply of jobs is confirmed by evidence that individuals with non-wage income whose wage earnings place them near the kink have a high propensity to track the kink despite having no incentive to do so (not reported).

In summary, the case study of teachers offers three lessons. First, firms tailor wage-hours offers to the average tax-induced preferences of their employees. Second, workers whose optimal earnings level is shared by many of their peers are more likely to respond to tax incentives and bunch at the kink. Finally, individuals who rely on the firm bunching mechanism to locate at the kink respond more to shifts in the top bracket cutoff than individual bunchers. We now show that each of these lessons generalizes to the full sample.

### 4.2.2 Coordinated Adjustment in other Occupations

We first test whether firms in other occupations systematically keep the peaks of their earnings distributions below the top tax bracket as do the teachers. We define the peak of the earnings distribution in each occupation-year cell in our data set as the DKr 5,000 wage earnings bin that has the largest density. Figure 19 shows the earnings distributions and our definition of the peaks for four selected occupation-years. Figure 20 plots the empirical distribution of the location of these peaks relative to the top tax bracket cutoff, shown by the vertical line at zero. The frequency of the peaks rises steadily until the top bracket cutoff, and then drops sharply after the threshold. There are 25 occupation-years with peaks in the band from 5,000 below to 5,000 above the kink. In contrast, there are only 4 occupation-years with peaks in the bands from 5,000 to 15,000 above the kink. We find this pattern of concentration of peaks below the top tax cutoff in each year of the sample as the bracket shifts (not reported), providing further evidence that occupation-level wage setting responds to tax incentives.[25] Note that these peaks are concentrated around the top tax cutoff. Across all occupations, earnings schedules appear to be tailored to the tax incentives of individuals who have little non-wage income and deductions, who are the most common in the population.

Next, we demonstrate that workers whose preferences are closer to the typical worker's preferences in their occupation are more likely to bunch at the kink. Consider two types of workers whose earnings place them in the top bracket. Both workers preferences are such that they would prefer to reduce their earnings to the top tax cutoff, because their marginal disutility of labor outweighs the marginal benefits of additional earnings given the high top

[25] Because the peaks are relatively well defined in most occupations, the peaks that are selected are not sensitive to the details of the algorithm used to identify peaks. In addition, a similar pattern is obtained when we select the 5 largest peaks from each occupation or identify local mass points by computing residuals from a regression on a polynomial.

tax rate. The first worker is in an occupation that has a large mass of workers near the kink, like the teachers. The second works in an occupation that has few workers earning incomes that place them near the top bracket. The first worker can rely on firm bunching in order to obtain a job that places him close to the kink, while the second worker must himself search for a job with lower hours and earnings. The model therefore predicts that the first worker is more likely to locate at the kink than the second. To test this prediction, Figure 21 plots the empirical distribution of earnings for two groups of workers: those in occupations with peaks less or more than DKr 25,000 from the top bracket cutoff. The excess mass around the kink for individuals in occupations with peaks near the kink is $b_f = 11.8$, compared with $b_f = 4.4$ in occupations with peaks far away from the kink.[26]

Figure 21 also shows the empirical distribution of earnings for those workers with adjustments larger than DKr 7,000 in occupations with peaks near the kink. Within this subset, $b_f = 4.4$, so there is actually no more bunching than for workers in non-bunching occupations. This is because occupations with peaks near the kink place only workers with 0 deductions at the kink. In order for these high-adjustment workers to bunch, they must individually find an unusual earnings package within the occupation. Bunching thus requires workers both to locate in an occupation with many potential bunchers and to have common tax preferences.

Finally, we investigate earnings dynamics around the top kink. We do so using the indicator defined in the previous section for wage earnings growth tracking the movement in the top tax cutoff over a two year period. Figure 22 shows that in occupations whose peaks are less than DKr 20,000 from the top bracket cutoff in the base year, the fraction of individuals for whom wage growth tracks the movement in the top bracket is maximized at the kink and falls rapidly thereafter. For these firm bunchers, the conditional probability of locating near the kink in year $t + 2$ given that one was near the kink in year $t$ is nearly 40%. In contrast, in occupations whose peaks are more than DKr 20,000 from the top bracket, wage growth for individuals near the kink is less likely to track the movement in the top bracket and is not noticeably different from wage growth patterns for others. The finding that individual bunchers do not respond to small movements in the top tax cutoff is consistent with the view

[26] These $b_f$ calculations are percentage measures normalized by the mass in the counterfactual distribution around the kink; hence, there is no mechanical reason to expect differences in $b_f$ across the two groups. We are simply comparing the probability that an individual whose earnings place him in the vicinity of the top kink locates at the kink itself.

that search costs attenuate behavioral responses to small tax changes, and provides further evidence against the view that our results are generated by non-constant elasticities. Finally, individuals who have more than DKr 7,000 of non-wage income but work in occupations whose modes are near the top tax cutoff still have an excess propensity to track the kink despite having no incentive to do so. This confirms that firm-level reoptimization (rather than individual job search) drives these dynamics. Though a good number of individuals do manage, through their own actions, to locate at the top tax kink, these individuals do not consistently move with the kink from year to year.

We conclude that the coordinated adjustments by firms play a central role in the effects of tax changes on the labor market in the long run. Changes in the supply of jobs in response to taxes may be particularly easy to detect in Denmark because of the prevalence of collective bargaining. Collective bargaining agreements, which cover 85% of the labor market, allow workers to coordinate wage and work schedules at a broad level. Although collective bargaining is less prevalent in other economies such as the U.S., technological constraints force coordination of work schedules and thereby produce hours constraints and other institutional rigidities in all labor markets. The general lesson to be drawn from the evidence here is that these institutional constraints are themselves endogenous to the tax regime. Although changes in taxes may not induce sharp, immediate responses in salaries set by firms as in the Danish case, they may induce changes in norms such as the length of workdays or job characteristics over time.

### 4.3 Prediction 3: Positive Covariance of Individual and Firm Bunching

The final prediction we test is that firm bunching and individual bunching should positively covary in the cross-section. We continue to examine firm bunching by looking at those individuals with significant deductions who, despite not facing a tax kink at the legislated threshold, nevertheless bunch at that point. We explored individual bunching above by selecting a set of individuals who have non-zero deductions or non-wage income, making it impossible for them to rely solely on their employer to reach the top kink. Part of the response we observe among workers with large adjustments is presumably due to manipulation of the adjustments themselves, though. For instance, individuals who happen to be in a job that pays DKr $20,000$ more than the top tax cutoff may elect to make a DKr $20,000$ contribution to their retirement account to avoid paying the top tax.

To investigate the covariance of firm and individual bunching, we must first isolate an empirical measure of individual bunching. There are two conceptual challenges in finding such a measure. Following equation 12 from our model, we must find a kink that applies to few enough individuals so that firms will essentially not cater to these workers' preferences when setting the distribution of job offers. Second, we must ensure that the bunching at this very uncommon kink is driven by labor supply choices and not the manipulation of deductions or non-wage income. Intuitively, we need an instrument for a worker's total adjustments to wage earnings.

To overcome these two empirical challenges, we take advantage of the cap on tax-deductible pension contributions, set at approximately DKr $33,000$. Figure 23a plots the distribution of deductions conditional on having deductions of at least DKr 20,000. There is a mass point in the distribution of deductions at precisely DKr $33,000$, demarcated by the green line.[27] These individuals reach the top tax bracket only when their wage earnings exceed the top tax cutoff by DKr $33,000$. This mass point in the distribution of pension contributions yields exogenous variation in deductions that can be used to identify the pure labor supply elasticity. What is more, the number of workers with deductions above DKr $20,000$ is a very small fraction of all workers, as shown in Figure 13. Therefore, those wishing to bunch at DKr $33,000$ above the top tax cutoff represent a very uncommon tax preference to which firms should not cater. In summary, the amount of bunching at this "pension kink" satisfies both of the requirements discussed above for a valid measure of individual bunching.

We implement our measure of individual bunching in Figure 23b, which plots the wage earnings distribution for the same group of individuals as in Figure 23a around the top tax cutoff. The wage earnings distribution reaches a global peak near the top tax cutoff as a result of firm bunching. But there is also a second peak in the wage earnings distribution at DKr $33,000$ above the legislated cutoff. This peak reflects the mass of individuals whose personal tax cutoffs are given by the green line. This figure shows that these individuals actively find higher-paying jobs that place them near the top tax cutoff given their pension

[27] The pension contribution cap increases slightly over years. The distributions of wage earnings have been recentered so that the pension cap falls at -33,000 in each year.

contributions.[28] This illustrates an interesting interaction between the incentives created by the tax system: the kink in taxes on retirement savings at $33,000$ produces a corresponding kink in labor earnings $33,000$ above the top tax cutoff.[29] We refer to this secondary tax kink as the "pension kink."

To confirm that firms do not significantly adjust the overall distribution of job offers in response to those at the pension kink, we next examine the distribution of wage earnings for workers with deductions close to zero, and therefore do not face any discrete change in incentives $33,000$ DKr above the legislated kink. Figure 23c plots this distribution; the distribution of wage earnings is nearly flat around the "pension kink," marked by the vertical line. Since there is no bunching in the distribution of wage earnings for these other workers, we know that workers saving the maximum of DKr $33,000$ in their pension are not disproportionately likely to initially draw a job at the "pension kink." Therefore, virtually the entire bunch at the "pension kink" represents individual bunching.

We now examine the covariance between firm and individual bunching in the cross-section, using this "pension-kink" measure of individual bunching. The two panels in Figure 24 plot the level of individual bunching (on the horizontal axis) against the amount of firm bunching (on the vertical axis) for a number of occupational and demographic subgroups. The correlation between these two series is visually striking and statistically significant. Figure 24a shows that pattern individual bunching across different occupations are mirrored in the patterns of firm bunching. For instance, professionals show the greatest amount of firm bunching and individual bunching across large occupational classes, while manual laborers demonstrate the least in each setting. Figure 24b shows a similar relationship across demographic groups. Women and the highly educated both demonstrate large amounts of firm and individual bunching, while a number of demographic groups, including men, single workers, and those with low education or experience tend to demonstrate the least amount of bunching.

This strong positive correlation between individual and firm bunching across the population demonstrates that job offer distributions, though set by firms, respond endogenously to worker

[28] To verify that this bunching in wage earnings is driven by individual selection of jobs rather than firm responses, we searched for and found no mass points in earnings distributions in this region for individuals without deductions.

[29] Analogously to what presented in Section 4.1.5, we find that small movements in top kink induce no changes in wage earnings for individual bunchers. This result rejects even more strongly the hypothesis that non-constant elasticity drives our finding of substantial clustering around the top tax.

preferences. In occupations where individual bunching is large, there must be a large mass of workers willing to incur the adjustment cost to reoptimize, reflecting a large labor supply elasticity. The evidence in Figure 24 shows that we should also expect a large amount of firm bunching, as firms offer a disproportionate share of jobs yielding wage earnings that place workers near the tax kink. But since firms have no innate preferences of their own that would drive this bunching in job offers, it must be that the individual preferences drive firm bunching as well.

Interestingly, much of the heterogeneity in bunching across demographic groups we initially documented in Figure 5 and show further in Figure 24b can also be traced to differences in search costs across occupations. Figures 25a and 25b show the distributions of taxable income for women ($b_f = 10.5\%$) and men ($b_f = 3.26\%$). This difference in observed elasticities could be driven either by preference heterogeneity or by differences in occupations across men and women. To distinguish between these two channels, we use the semi-parametric propensity score reweighting technique proposed by DiNardo et al. (1996) to control for the differential distribution of occupations between men and women. In Figure 25c, we reweight the sample of men to match the observed distribution of occupations for women, effectively placing more weight on those men who work in female-dominated occupations. This reweighting increases the excess mass observed among men to $b_f = 8.55\%$. This simple occupational reweighting explains approximately 75% of the observed difference in excess mass at the kink between men and women, suggesting that differences in occupational characteristics drive most of the heterogeneity in tax responsiveness across men and women. Consistent with this interpretation, we find that women work in occupations that have thicker labor markets. On average, women work in occupation-regions that have 43,674 workers, whereas the average occupation-region size for men is 29,091. This evidence shows that commonly accepted intuitions about labor supply elasticities – such as higher elasticities for women – may be driven by variation in frictions rather than fundamental heterogeneity in the structural elasticities that determine long-run behavior.

Thus far, we have focused on variation in search costs across wage earners. The most flexible group of workers in the economy are the self-employed, who face weaker hours constraints and do not need to search for a different job to change their earnings. Self employed individuals also have the greatest flexibility to change reported incomes, either by transferring income across years to minimize tax burdens, or by under-reporting their taxable incomes. Figure 26a shows the distribution of taxable income around the top tax cutoff for individuals who report positive self-employment income. The estimated excess mass is $b_f = 150\%$, dwarfing the estimated excess for wage earners by an order-of-magnitude. Figure 26b shows the distribution of taxable income around the largest middle tax cutoff, where net of tax wages fall by 10%, for self-employed individuals. Unlike in the case of wage earners, there is substantial bunching at each of the smaller kinks in the tax schedule for the self-employed.

We suspect that most of the bunching among the self-employed is not driven by changes in labor supply but rather by reporting effects and intertemporal substitution.[30] Indeed, LeMaire and Schjerning (2007) identify the same excess mass in the distribution of reported taxable earnings for the self-employed but find no corresponding bunching in the distribution of profits from the self-employment enterprise. Furthermore, Kleven et al. (2009) uncover substantial tax evasion among the self-employed. More specifically for our purposes, they find that at least 40% of all bunching at the top kink is driven by illegal tax evasion. Nevertheless, we can conclude that both small and large tax variation induces substantial changes in behavior on flexible margins (reporting or self-employed earnings), but only large tax changes induce responses on margins that are harder to adjust (wage earnings).

## 5 Calibration

The model and evidence above suggests that observed elasticities are the result not only of the structural elasticity parameter, as in the frictionless case, but also of individual adjustment costs and the endogenous firm responses. Unfortunately, the model is only identified non-parametrically for an infinitely large tax change with infinitessimal scope. This identification at infinity argument relies on two pieces of intuition. First, if a tax change were infinitely large, then all workers would respond. Individual bunching would reflect the full responsive of preferences. Second, since the tax kink applies to very few people, firms would not alter the equilibrium distribution of jobs at all, and so there would be no firm bunching.

Although the identification strategy described above is useful in guiding intuition, it cannot be implement in practice with the finite tax changes we have in our data. We therefore pursue

[30] The Danish tax code allows the self-employed to shift some income across years legally.

an alternative partial identification approach here. We show that the observed elasticities at the middle and top kinks together place a lower bound on the structural elasticity if one restricts attention to adjustment costs below some threshold. To understand the forces that allow us to obtain such a bound, first consider the impact of the adjustment cost scale $\phi$ and the structural elasticity on the observed amount of bunching. Figure 27a plots the relationship between predicted bunching at the top kink (vertical axis) against the structural elasticity parameter for three different values of $\phi$ (and holding $\gamma$ fixed at 1.1). When $\phi = 0$, the amount of bunching is a linearly increasing function of the structural elasticity. As the elasticity increases, the fraction of workers who optimally relocate to the kink is increasing in the underlying elasticity. Once we introduce frictions, though, this relationship becomes an inverted U-shape. As the structural elasticity increases, more workers still share the kink as an optimal labor supply choice. But larger structural elasticities also imply that workers have a less concave utility function, and so they suffer smaller losses as a result of deviations from optimal labor supply. Chetty (2009a) uses a Taylor approximation to show that the utility cost of setting labor supply at $l$ instead of $l^*$ is approximately

$$u(l^*) - u(l) = -\frac{1}{2}\frac{1}{\varepsilon}wl^*(\Delta \log l)^2 \tag{16}$$

which implies directly that $\frac{\partial u(l^*)-u(l)}{\partial \varepsilon} < 0$. As a result, fewer workers choose to reoptimize, which decreases the observed amount of bunching. The first force, which increases the amount of bunching, dominates for low elasticities, while the latter force drives the amount of bunching back down at larger ones. Furthermore, as the scale of the adjustment costs increase, the peak of the inverted U-shape decreases. In Figure 27a, for instance, we plot the inverted-U relationship for $\phi = 10\%$ and $\phi = 8\%$ of consumption. The observed amount of bunching begins to decrease at $\varepsilon = 0.2$ for the larger adjustment cost but not until $\varepsilon = 0.25$ for the smaller adjustment cost.

Now consider fitting these predictions to the data. The horizontal dashed line at $bf = 14.1\%$ marks the observed level of bunching at the top kink. Each inverted U-shape in Figure 27a crosses the observed $bf = 14.1$ twice, implying that two structural elasticities each replicate bunching at the top kink: one very small structural elasticity and one much larger value. We discuss the very different intuition between each of these "crossing values" now in turn. Intuitively, the small structural elasticity corresponds to a case in which workers suffer large losses from suboptimal labor supply, as in equation 16, and thus most workers pay to reoptimize. The small structural elasticity implies that relatively few workers share an optimal point at the kink, however, and therefore the predicted amount of bunching is small. This smaller implied elasticity cannot be consistent with our findings, however, since the estimated elasticity will not vary with the size of the underlying tax change. This would not produce the observed upward-sloping relationship between the estimated elasticity and the size of the kink. To demonstrate this fact, Figure 27b plots these inverted-U relationships for the $\phi = 8\%$ case in Figure 27a but for both the middle and the top kink. The amounts of bunching observed are once again marked with horizontal lines. In order to be consistent with the data, the generated curves must each cross the lines representing the respective observed elasticity at the same value of $\varepsilon$ the two lower crossing values are quite different, while the upper crossing value is identical

The intuition behind the larger structural elasticity consistent with observed bunching at the top kink is very different. In this case many workers wish to relocate to the kink, but their relatively elastic preferences make most of them unwilling to pay the adjustment cost to do so. As one increases the scale of the adjustment costs, this larger implied elasticity shrinks substantially, since more inelastic workers who suffer larger utility losses from suboptimal labor supply become unwilling to pay to relocate. Counterintuitively, larger adjustment costs therefore imply a *smaller* structural elasticity.

To place bounds on the structural elasticity consistent with our findings, we must now quantitatively answer the question: how large must adjustment costs be to imply a structural elasticity of a given size? We therefore calibrate our numerical model from Section 2.5 using the observed bunching at kinkpoints and then estimate the value of the structural elasticity parameter most consistent with the data. We use the distribution of income *away* from the kinks to generate the underlying distribution of the taste parameter $\alpha_i$; therefore, our eventual elasticity estimate uses only the moments that are bunching at kinkpoints for fit. We assume the precision cost function is $c(q) = 1 + q^{1+\gamma}$. This simple function form includes both a fixed cost of search, as well as a increasing marginal cost to precision. We also scale $\phi$ as a fraction of optimal consumption for each alpha-type so that we mechanically create neither an increasing nor decreasing propensity to search as a worker's taste for leisure increases.

Therefore if $\phi = 0.1$, then workers would face a cost of 10% of optimal consumption to begin to search and then another 10% to search perfectly. We assume that the diffuse non-directed search follows a normal distribution with standard deviation matching the year-on-year change in income, which we calibrate from the data as DKr $40,000$. We solve for the equilibrium hours offer distribution by iteration (a la Blackwell's theorem) and then check that we have found the unique equilibrium by starting the iteration from different points. To approximate non-wage income in the data, we add a random draw from a normal distribution with SD equal to DKr $5,000$ to each worker's labor income. As in the empirical work, we estimate the counterfactual polynomial as a seven-degree polynomial on the equilibrium distribution within DKr $50,000$ on either side of the tax kink. We set the fraction of workers that share each kink to $\zeta = 0.6$, to match the data. This leaves three free parameters to estimate — the structural elasticity $\varepsilon$ and the elasticity of search costs $\gamma$, and the adjustment cost scale parameter $\phi$. Since we have only two moments from the data — the bunching fraction at the top kink ($bf = 14.1\%$) and at the middle kinks ($bf = 1.2\%$) — we are not point-identified. We therefore calculate the implied structural elasticity given a wide range of plausible value of $\phi$.

Figure 27c plots the implied structural elasticity $\varepsilon$ (on the vertical axis) for each potential value of the scale of adjustment costs $\phi$ (on the horizontal axis). Because of the declining cost of suboptimal choices for higher elasticities discussed above, this relationship slopes downward. Even for large adjustment costs, though, the implied elasticities fall no smaller than 0.3. We consider $\phi = 10\%$ of consumption as an upper bound on the adjustment cost. This would imply that workers must pay 20% of consumption to select their optimal job perfectly. This magnitude of adjustment costs would also imply that, on average, workers leave 8% of consumption on the table by making sub-optimal labor supply choices in each year. This upper bound on $\phi$ implies a lower bound of $\varepsilon \geq 0.28$. Even at this bound, the estimated elasticity is an order of magnitude larger than the observed elasticity of $\hat{\varepsilon} = 0.029$ at the top kink.

# 6 Conclusion

This paper has investigated the impacts of search costs and hours constraints in the labor market on earnings responses to income taxes. A stylized model of frictions and labor supply predicts that observed elasticities will be larger when (a) the tax change is larger, and (b) the change affects incentives for a larger group of individuals. The model also predicts a strong positive correlation between individual and firm responses reflecting the endogenous determination of firm wage-hours packages. Using administrative tax records for the population of Denmark, we document evidence supporting each of these predictions. Large tax changes at kink points in the tax schedule generate sharp bunching, whereas smaller kinks and tax reforms generate virtually no response. The bunching response is driven by a combination of workers selecting job offers that allow them to locate near the large kink given their deductions and other income, and firms tailoring jobs to allow individuals with common tax incentives to locate at the kink. The firms' supply of jobs response is significantly greater in occupations where a large mass of individuals have local incentives distorted by the kink, consistent with the prediction that tax distortions that apply to a wide group of individuals generate greater elasticities. Firm bunching and individual bunching covary across different subgroups of the population, since worker preferences drive the variation across groups.

We conclude that failing to account for search frictions and institutional constraints can lead to substantially attenuated estimates of the long-run impact of taxes on labor supply. When we implement the widely applied method of Gruber and Saez (2009) for estimating taxable income elasticities, we precisely estimate an effect of taxes near zero. Our analysis demonstrates that examining the same data in a manner that takes frictions into account uncovers earnings responses to income taxation that are an order of magnitude higher.

The empirical results suggest that income taxes could potentially have larger efficiency costs than implied by previous elasticity estimates in the microeconometric literature. However, additional theoretical work is necessary to characterize the welfare consequences of tax policies in a model with frictions and endogenous hours constraints because workers are not at their optimal hours choices. For example, tax policies could potentially have larger efficiency costs than implied by the structural preference elasticity if they force inelastic high earners to coordinate hours with more elastic lower earners. Another example concerns the prediction that it is optimal to levy higher tax rates on men than women because they are less elastic (Boskin and Sheshinski 1983, Alesina et al. 2007, Kleven et al. 2009). If the difference in observed elasticities is caused by heterogeneity in frictions rather than tastes, as the evidence above suggests, taxes may distort the behavior of men as much as women in the long run. In future work, it would be interesting to analyze the efficiency costs of taxes and optimal tax policy by building on the simple model in Section 2.

Our results suggest that the policy-relevant elasticity parameter, in many settings, is much larger than microeconometric estimates. Even the largest tax changes may not be sufficiently large to overcome frictions, and our model implies that the true elasticity can only be recovered non-parametrically in the limit as tax changes become infinitely large. An interesting direction for further work would be to identify the true elasticity by imposing more structure on the model. Furthermore, our lower bound of 0.28 on the elasticity parameter relies only on adjustment that *could* occur in the short run but does not. Other channels of adjustment, such as changes in the stocks of human or physical capital, take more time. Theoretically, these effects could amplify or negate the responses we document, and so further work is required to link our estimates to directly policy-relevant numbers.

Finally, our analysis serves as a caution in interpreting the empirical estimates obtained from quasi-experimental studies in the modern applied microeconomics literature. Although focusing on small experiments is extremely useful for identification of causal effects, the behavioral responses identified by such studies are particularly susceptible to attenuation by frictions. Credibly identifying long-run responses in a manner that accounts for frictions would be valuable not just in the context of labor supply but also in many other settings.

# 7 References


1. Altonji, Joseph G. and Christina H. Paxson (1988). "Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-Offs," *Journal of Labor Economics*, 6(2): 254-76.

2. Altonji, Joseph G. and Christina H. Paxson (1992). "Labor Supply, Hours Constraints, and Job Mobility." *Journal of Human Resources*, 12: 269-310.

3. Alesina, Alberto F., Edward L. Glaeser, and Bruce Sacerdote (2005). "Work and Leisure in the U.S. and Europe: Why So Different?" *NBER Maco Annual*, pp. 1-99.

4. Alesina, Alberto F., Andrea Ichino, and Loukas Karabarbounis (2007). "Gender-Based Taxation and the Division of Family Chores." NBER Working Paper 13638.

5. Blundell, Richard, Mike Brewer, Marco Francesconi (2008). "Job Changes and Hours Changes: Understanding the Path of Labor Supply Adjustment." *Journal of Labor Economics*, 26(3): 421-453.

6. Blundell, Richard, Alan Duncan, and Costas Meghir (1998). "Estimating Labor Supply Responses Using Tax Reforms." *Econometrica*, 66: 827-861.

7. Blundell, Richard, and Thomas MaCurdy (1999). "Labor Supply: A Review of Alternative Approaches." *Handbook of Labor Economics*, Vol 3A, Ashenfelter and Card eds., pp. 1559-1695.

8. Boskin, Michael J. and Eytan Sheshinski (1983). "Optimal Tax Treatment of the Family: Married Couples," *Journal of Public Economics* 20, pp. 281-297.

9. Burtless, Gary and Jerry A. Hausman (1978). "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." *Journal of Political Economy*, 86: 1103-1130.

10. Chetty, Raj (2009a). "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." NBER Working Paper No. 15616.

11. Chetty, Raj (2009b). "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance." Forthcoming, *American Economic Journal — Economic Policy.*

12. Chetty, Raj, Adam Looney, and Kory Kroft (2009). "Salience and Taxation: Theory and Evidence." *American Economic Review* 99(4): 1145-1177.

13. Chetty, Raj and Emmanuel Saez (2009). "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients." NBER Working Paper No. 14836.

14. Chetty, Raj, Anders Frederiksen, John Friedman, Tore Olsen, and Luigi Pistaferri (2010). "What Determines Knowledge of Tax Policies? Evidence from Denmark." Harvard University Working Paper.

15. Cogan, John (1981), "Fixed Costs and Labor Supply." *Econometrica*, Vol. 49(4): 945-63.

16. Davis, Stephen J., and Magnus Henrekson (2005). "Wage-Setting Institutions as Industrial Policy." *Labour Economics*, Vol 12, pp. 345-377.

17. Dickens, William T. and Shelly J. Lundberg (1993). "Hours Restrictions and Labor Supply." *International Economic Review*, 34(1): 169-92.

18. DiNardo, John, Nicole M. Fortin, Thomas Lemieux (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica*, Vol 64, pp. 1001-1044.

19. Eissa, Nada and Jeffrey B. Liebman (1996). "Labor Supply Response to the Earned Income Tax Credit." *Quarterly Journal of Economics*, 111(2): 605-37.

20. Feldstein, Martin (1995). "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of TRA 1986." *Journal of Political Economy*, Vol. 103(3), pp. 905-926.

21. Gelber, Alex (2008). "Taxation and Family Labor Supply." University of Pennsylvania Working Paper.

22. Goolsbee, Austin (2000). "What Happens When You Tax the Rich? Evidence from Executive Compensation." *Journal of Political Economy*, Vol 108, pp. 352-378.

23. Grossman, Sanford J. and Guy Laroque (1990). "Asset Pricing and Optimal Portfolio Chioce in the Presence of Illiquid Durable Consumption Goods." *Econometrica*, Vol 58, pp. 25-51.

24. Gruber, Jonathan and Emmanuel Saez (2002). "The Elasticity of Taxable Income: Evidence and Implications" *Journal of Public Economics*, Vol. 84(1), pp. 1-32.

25. Ham, J.C. (1982). "Estimation of a Labour Supply Model with Censoring Due to Unemployment and Underemployment." *Review of Economic Studies*, Vol. 49(3), pp. 335-354.

26. Hausman, Jerry A. (1981). "Labor Supply." In *How Taxes Affect Economic Behavior*, eds. Henry Aaron and Joseph Pechman.

27. Hausman, Jerry A. (1983). "Tax and Labor Supply." In *Handbook of Public Economics*, eds, Alan Auerbach and Martin Feldstein.

28. Heckman, James J. "What Has Been Learned About Labor Supply in the Past Twenty Years?" *American Economic Review*, 83(2), pp. 116-121.

29. Kleven, Henrik, Martin Knudsen, Claus T. Kreiner, Søren Pedersen, and Emmanuel Saez (2009). "An Experimental Evaluation of Tax Evasion and Tax Enforcement in Demark." UC Berkeley Working Paper.

30. Kleven, Henrik, Claus Kreiner, and Emmanuel Saez (2009). "The Optimal Income Taxation of Couples." *Econometrica*, Vol 77, pp. 537-560.

31. Kleven, Henrik and Esben A. Schultz (2009). "Estimating the Elasticity of Reported Income Using a Large Panel of Danish Taxpayers." LSE Working Paper.

32. Kopchuk, Wojciech, and Joel B. Slemrod (2002). "The Optimal Elasticity of Taxable Income." *Journal of Public Economics*, Vol 84, pp. 91-112.

33. Lange, Katherine, Lars Haagen Pedersen and Peter Birch Sorensen (1999). "The Danish Tax Reform Act of 1993: Effects on the Macroeconomy and on Intergenerational Welfare." In *Macroeconomic Perspectives on the Danish Economy*, Torben M. Andersen and Svend E. Hougaard Jensen (eds.), pp. 134-161.

34. LeMaire, Daniel, and Bertel Schjerning (2007). "Earnings, Uncertainty, and the Self-Employment Choice." CBER Discussion Paper DP 2007-04.

35. Lindbeck, Assar (1995)."Welfare State Disincentives with Endogenous Habits and Norms." *The Scandinavian Journal of Economics*, Vol. 97(4), pp. 477-494.

36. MaCurdy, Thomas (1981). "An Empirical Model of Labor Supply in a Life-Cycle Setting." *Journal of Political Economy*, Vol 89, pp. 1059-1085.

37. MaCurdy, Thomas, David Green and Harry Paarsche (1990). "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply." *The Journal of Human Resources*, Vol 25, pp. 415-490.

38. Meghir, Costas, and David Phillips (2008). "Labour Supply and Taxes." IZA Discussion Paper #3405.

39. Meyer, Bruce and Dan Rosenbaum (2001). "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics*, Vol. 116(3), pp 1063-1114.

40. Mirrlees, James A. (1971). "An Exploration in the Theory of Optimum Income Taxation." *The Review of Economic Studies*, Vol 38(2), pp. 175-208.

41. Moffitt, Robert A. (1990). "The Econometrics of Kinked Budget Constraints." *Journal of Economic Perspectives*, Vol 4, pp. 119-139.

42. Moffitt, Robert A., and Mark Wilhelm (2000). "Taxation and the Labor Supply Decision of the Affluent." In *Does Atlas Shrug? Economic Consequence of Taxing the Rich*, ed. Joel Slemrod, pp. 193-234.

43. OECD (2009). Factbook 2009: Economic, Environmental and Social Statistics. http://puck.sourceoecd.org/vl=13806301/cl=14/nw=1/rpsv/factbook2009/06/03/02/index.htm

44. Ohanian, Lee, Andrea Raffo, and Richard Rogerson (2008). "Long-Term Changes in Labor Supply and Taxes: Evidence from OECD Countries, 1956-2004." Forthcoming, *Journal of Monetary Economics*.

45. Prescott, Edward C. (2004). "Why Fo Americans Work So Much More Than Europeans?" *Federal Reserve BAnk of Minneapolis Quarterly Review*, Vol 28, pp. 2-13.

46. Rogerson, Richard (1988). "Indivisible Labor, Lotteries and Equilibrium." *Journal of Monetary Economics* 21, pp 3–16.

47. Rogerson, Richard and Johanna Wallenius (2008). "Micro and Macro Elasticities in a Life Cycle Model with Taxes." Forthcoming, *Journal of Economic Theory*.

48. Rosen, Harvey S. (1976). "Taxes in a Labor Supply Model with Joint Wage-Hours Determination," *Econometrica*, 44(3): 485-507.

49. Saez, Emmanuel (2009). "Do Taxpayers Bunch at Kink Points?" Forthcoming. *American Economic Journal: Economic Policy*.

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

51. Saez, Emmanuel (2004). "Reported Incomes and Marginal Tax Rates, 1960-2000: Evidence and Policy Implications." *Tax Policy and the Economy*, Vol 18, ed. Jim Poterba.

52. Saez, Emmanuel, Joel B. Slemrod, and Seth H. Giertz (2009). "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." Forthcoming, *Journal of Economic Literature*.

53. Sammartino, Frank, and David Weiner (1997). "Recent Evidence on Taxpayers' Response to the Rate Increases of the 1990s." *National Tax Journal*, Vol 50, pp. 683-705.

54. Slemrod, Joel B. (1998). "Methodological Issues in Measuring and Intepreting Taxable Income Elasticities." *National Tax Journal*, Vol. 51(4), pp. 773-788.

55. Slemrod, Joel and Shlomo Yitzhaki (2002). "Tax avoidance, evasion, and administration." In *Handbook of Public Economics*, Alan Auerbach and Martin Feldstein (eds.), pp. 1423-1470.

## Appendix A: Theoretical Derivations

*Robustness of Proposition 1 to Non-Isoelastic Utility.* Consider a model without adjustment costs where the elasticity of labor supply depends arbitrarily on the net-of-tax rate so that

$$\varepsilon\left(w_i\left(1-\tau\right)\right)=\frac{\partial\ln h^*}{\partial\ln\left(w\left(1-\tau\right)\right)}>0.$$

and therefore

$$\ln h^*\left(w_i\left(1-\tau\right)\right)=\int_0^{w_i(1-\tau)}\varepsilon\left(\ln s\right)ds.$$

For ease of exposition, we suppress the individual taste heterogeneity in this section by setting $\alpha_i=1$ for all workers. Suppose then that there is a tax kink at earnings level $K$ at which the tax rate rises from $\tau$ to $\tau'$. Workers choose to bunch at this kink if and only if $w_i h^*\left(w_i\left(1-\tau\right)\right)\geq K\geq w_i h^*\left(w_i\left(1-\tau'\right)\right)$ which will hold over a range $\left[\hat{w}\left(\tau\right),\hat{w}\left(\tau'\right)\right]$.where $\hat{w}\left(\tau\right)=w \mid \; wh^*\left(w\left(1-\tau\right)\right)=K$ is a weakly increasing function of $\tau$. The amount of bunching at such a tax kink is $B\left(\tau,\tau'\right)=\int_{\hat{w}(\tau)}^{\hat{w}(\tau')}dF\left(w_i\right)$. Therefore, for any tax rates $\tau_1<\tau_2<\tau_3$ a model without adjustment costs predicts the relationship

$$B\left(\tau_1,\tau_3\right)=B\left(\tau_1,\tau_2\right)+B\left(\tau_2,\tau_3\right)$$

so that the amount of bunching created from two smaller kinks is exactly equal to the bunching created at one larger kink.

Suppose that we now consider the same three kinks in a model with adjustment costs. From above, we know that the amount of bunching is now only $\hat{B}=\theta B<B$, where the proportion of workers $\theta$ paying the adjustment cost to locate optimally is increasing in the size of the tax kink. Therefore the adjustment costs model instead predicts

$$\hat{B}\left(\tau_1,\tau_3\right)>\hat{B}\left(\tau_1,\tau_2\right)+\hat{B}\left(\tau_2,\tau_3\right).$$

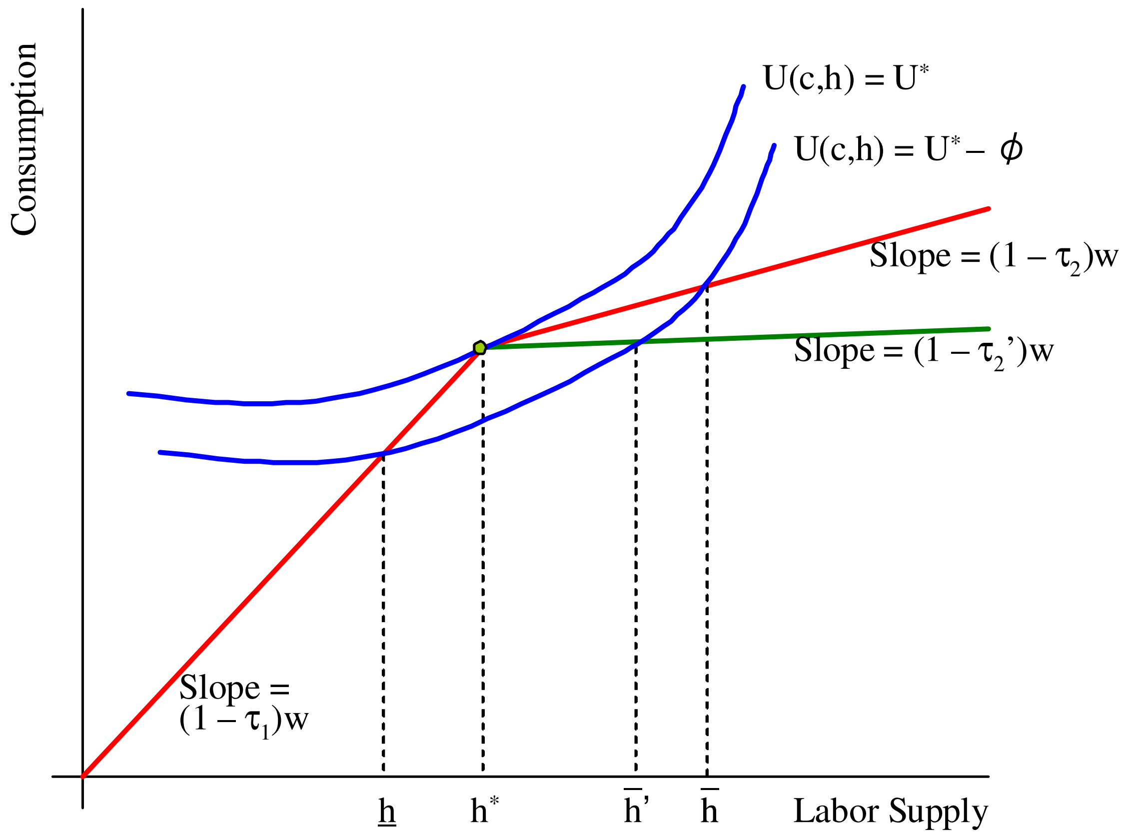
Since the amount of bunching increases faster than proportionately with the size of the tax kink, two smaller kinks should produce less bunching than one larger kink.

## Appendix B: Numerical Simulations and Calibration

The numerical model uses a discrete grid of alpha-types to approximate the population of workers, as well as discrete grid of available hours that match the set of optimal hours locations for these alpha-types. We choose the size of the tax kink so that $\{\underline{\alpha}, \bar{\alpha}\}$ are both among the alpha types. By varying the number of alpha-types in between these two endpoints, we vary the size of the tax kink. Since the utility function produces a log-linear labor supply function, the grid is evenly spaced in log-points, so that the distance between each of the 401 points is 1.005. We choose the distribution of alphas to match the distribution of earnings in the Danish population away from the kink. We assume the precision cost function is $c(q) = 1 + q^{1+\gamma}$. This simple function form includes both a fixed cost of search, as well as a marginal cost to precision. We also scale $\phi$ as a fraction of optimized utility for each alpha-type so that we mechanically create neither an increasing nor decreasing propensity to search as a worker's taste for leisure increases. Therefore, if for instance $\phi = 0.1$, then workers would face a cost of 10% of optimal utility to begin to search and then another 10% to search perfectly. We assume that the diffuse non-directed search follows a normal distribution with standard deviation matching the year-on-year change in income, which we calibrate from the data as DKr $40,000$. We solve for the equilibrium hours offer distribution by iteration (a la Blackwell's theorem), and then check that we have found the unique equilibrium by starting the iteration from different points. To approximate non-wage income in the data, we add a random draw from a normal distribution with SD equal to DKr $5,000$ to each worker's labor income. As in the empirical work, we estimate the counterfactual polynomial as a seven-degree polynomial on the equilibrium distribution within DKr $50,000$ on either side of the tax kink.

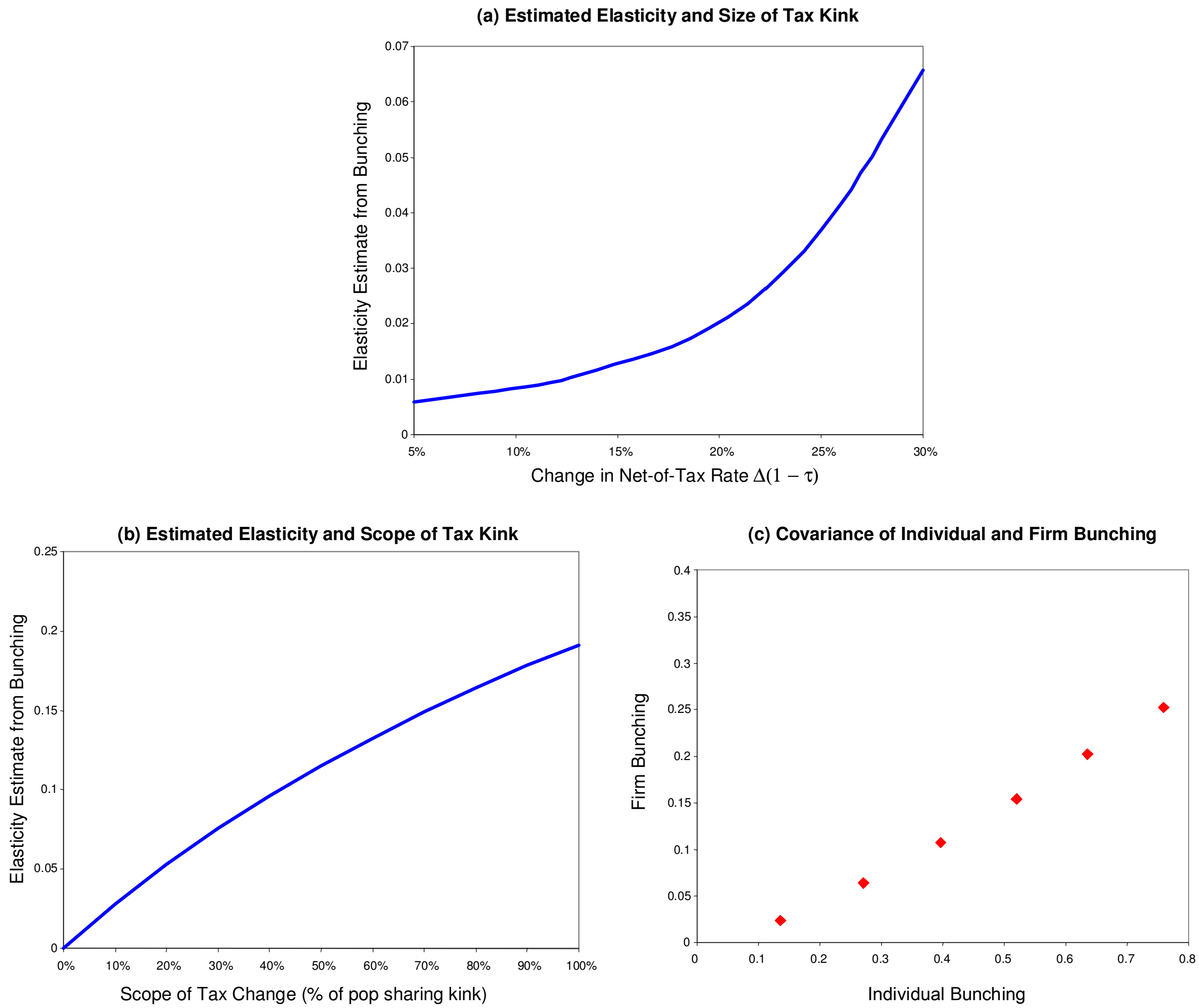
In Figure 2a, we use $\varepsilon = 0.4$, $\gamma = 0.05$, $\phi_i = 0.11$ for all workers, and $\zeta = 1$. We vary the number of alpha-types bunching from 4 to 24 to generate the variation in the log change in the net-of-tax rate from 3% to 30%. In Figure 2b, we use $\varepsilon = 0.4$, $\gamma = 0.05$, $\phi_i = 0.09$ for all workers and $\Delta \log(ntr) = 15\%$ (implying that 13 alpha-types find it optimal to bunch). In Figure 2c, we use $\gamma = 0.05$, $\phi_i = 0.09$ for all workers, $\zeta = 0.8$, and $\Delta \log(ntr) = 20\%$ (implying that $40\varepsilon + 1$ alpha-types find it optimal to locate at the kink for each different structural elasticity $\varepsilon$). We match the empirical distribution of income around the top kink in the Danish data in each figure.

FIGURE 1
Bunching at Kinks with Search Costs



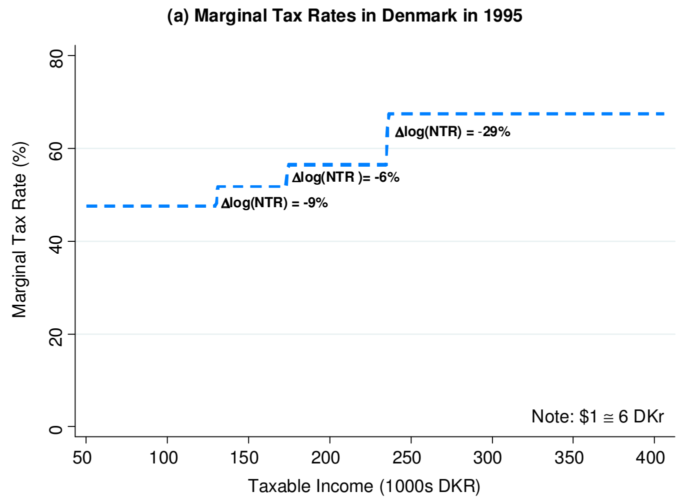

Notes: This figure illustrates the search problem faced by workers choosing labor supply. The two-bracket tax system creates the kinked budget set shown in red. The worker's indifference curves in $(c,h)$ space are illustrated by the blue curves. This worker's optimal labor supply is to set $h^*$ at the tax bracket cutoff. The lower indifference curve shows the optimal utility minus the search cost $\phi$. At any point between the two indifference curves—that is, any labor supply between $\bar{h}$ and $\underline{h}$— the worker will not pay $\phi$ to reoptimize. As the tax change at the bracket cutoff increases in magnitude (shown by the green budget line), the kink in the budget set becomes sharper, and the inaction region shrinks.

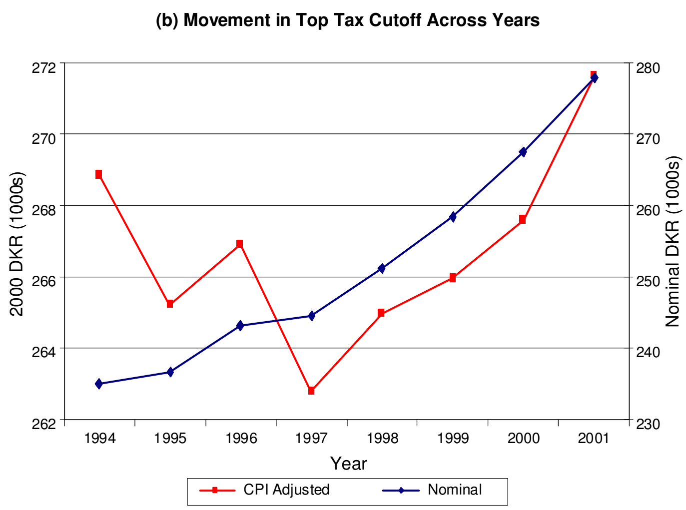
FIGURE 2
Three Testable Predictions: Numerical Simulations

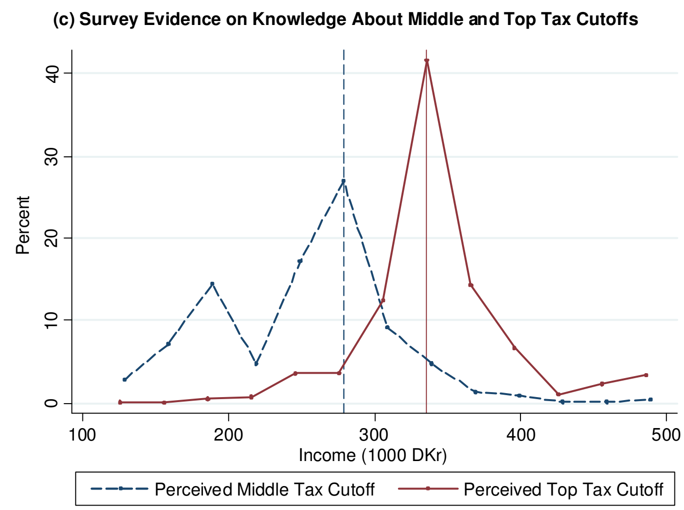



Notes: These figures present numerical simulations of elasticities estimated from the amount of bunching at kinks in the general model. Panel A shows the relationship between the estimated elasticity and the change in the net-of-tax rate at the kink. Panel B shows the relationship between the estimated elasticity and the scope of the tax change, defined as the fraction of the population affected by the kink. Panel C demonstrates how individual bunching and firm bunching covary as the structural elasticity increases. See Appendix B for the parameter values and functional form assumptions used to draw these figures.

# FIGURE 3
## The Danish Income Tax System

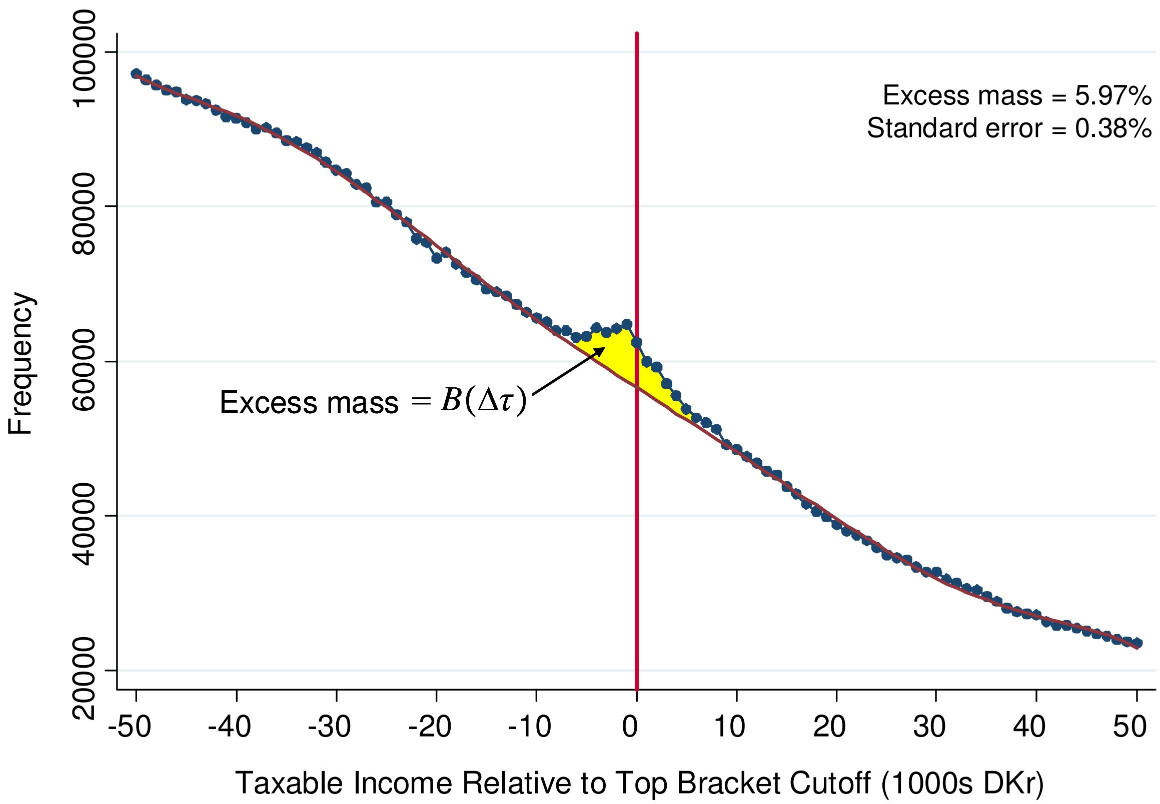


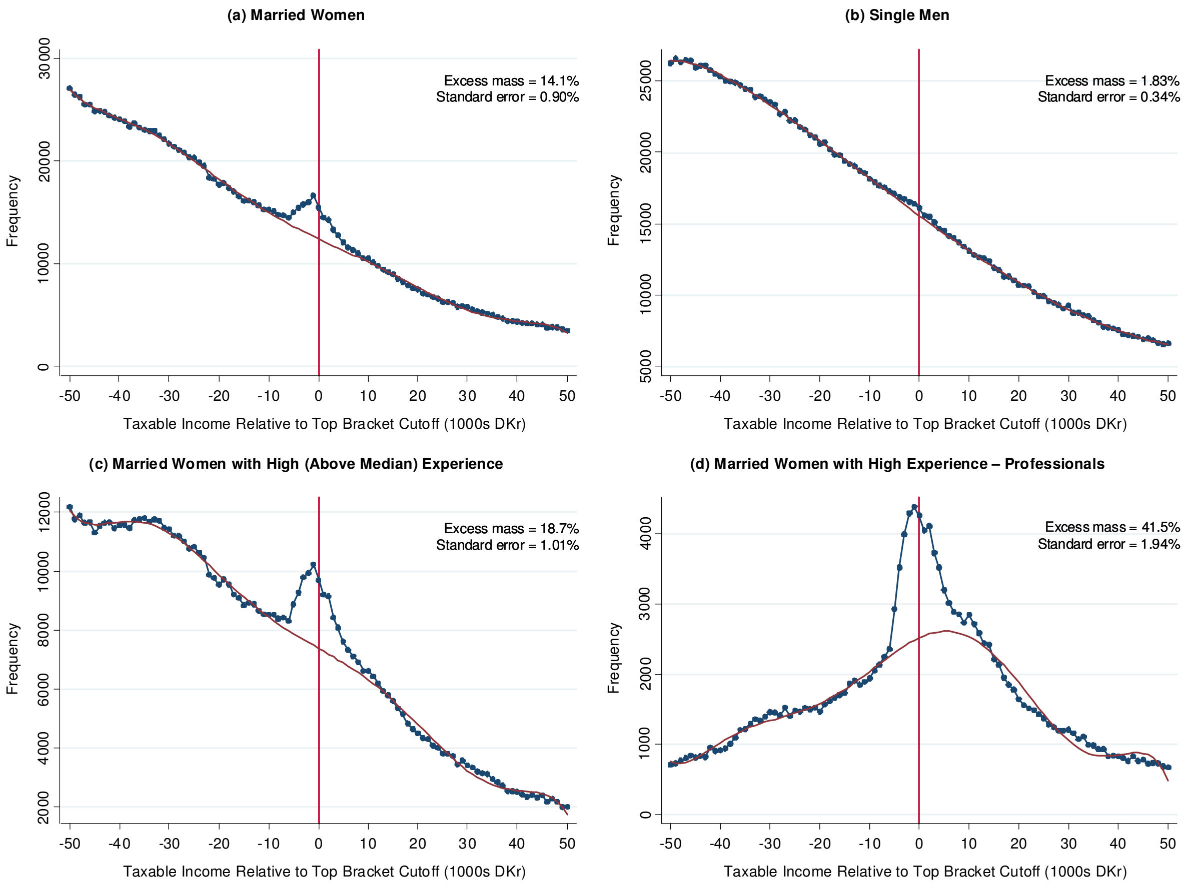

Notes: Panel (a) plots the marginal tax rate in 1995 for Danish earners by income level, including the state tax, region tax, and average municipal tax. Panel (b) plots the statutory earnings level above which earners must pay the top bracket state tax. The blue diamond dot series, plotted on the right *y*-axis, shows the nominal earnings level; the red square dot series, plotted on the left *y*-axis, shows the level in real 2000 DKr. Panel (c) plots the distribution of responses from 3,299 participants in a pension organization internet survey who were asked at which income levels the top tax bracket and middle tax bracket begin in the 2008 Danish tax code. The fraction of responses for the top tax (solid line) and middle tax (dashed line) cutoffs are shown in DKr 30,000 wide bins centered on the true interval, so that the mode of each distribution represents the fraction of people who knew within DKr 15,000 the correct tax bracket cutoff point.

FIGURE 4
Income Distribution around Top Tax Cutoff for Wage Earners



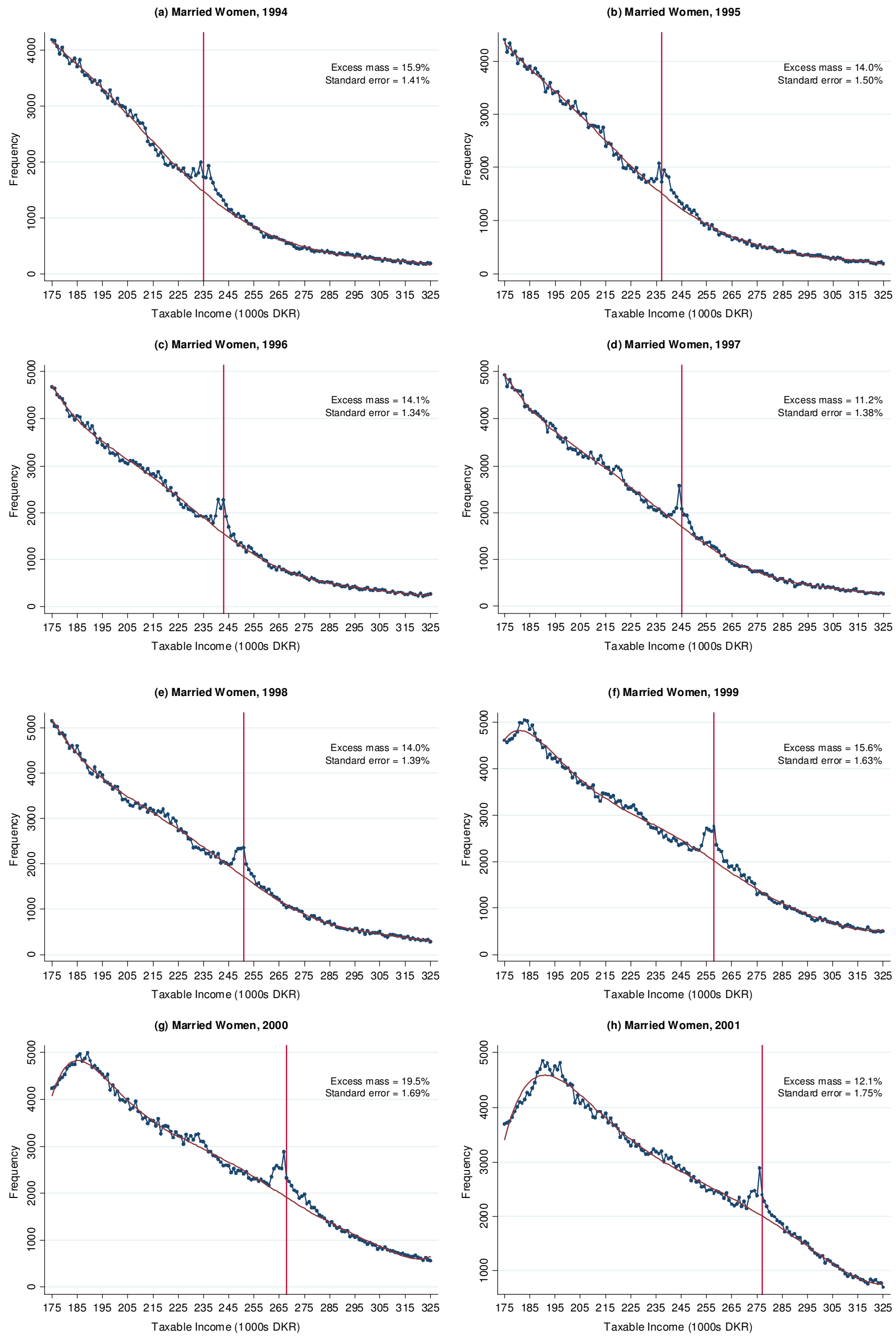

Notes: This figure shows the taxable income distribution around the top tax bracket cutoff. The series with dots plots a histogram of taxable income (in this case, the top tax base) for wage earners between 1994-2001, relative to the level above which earners must pay the top tax in each year. Each point shows the number of observations in the dataset in a DKr 1,000 bin. The solid line beneath the empirical distribution is a seventh-degree polynomial fitted to the empirical distribution excluding the datapoints DKr 7,000 or fewer from the cutoff. The shaded region between the two lines is our measurement of the excess mass at the top bracket cutoff, which is 5.97% of the area within the same domain on the $x$-axis beneath the counterfactual distribution. The standard error for the excess mass is obtained from the standard errors on the regression coefficient estimates based on the tightness of fit of the polynomial.

# FIGURE 5
## Heterogeneity in Bunching at the Top Tax Cutoff




Notes: These four figures plot the empirical distributions of taxable income, as well as the counterfactual distributions and excess masses, for wage earners between 1994-2001 who are (a) Married Women, (b) Single Men, (c) Married Women with Experience Greater than 18 Years, and (d) Married Women with Experience Greater than 18 Years who have Danish ISCO code in the 2000s (professionals). For the methodology for calculating the counterfactual distributions, see the notes to Figure 4.

# FIGURE 6
## Income Distributions for Married Women 1994-2001

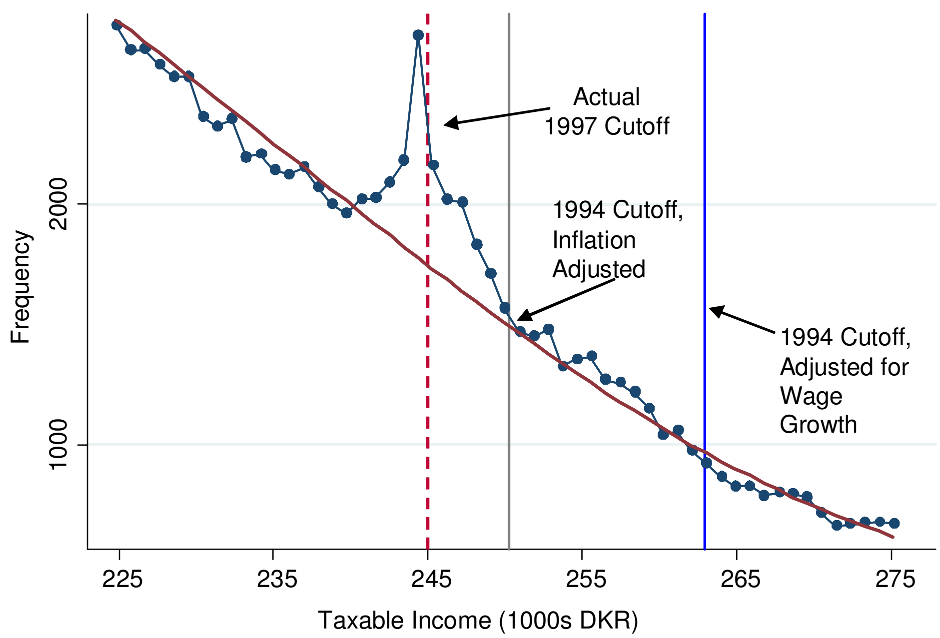



Notes: These figures plot the empirical distribution of taxable income, as well as the counterfactual distributions and excess masses, for married female wage earners in each year from 1994-2001. The solid vertical lines mark the top tax bracket cutoff, in nominal DKr, in each year. For the methodology for calculating the counterfactual distributions, see the notes to Figure 4.

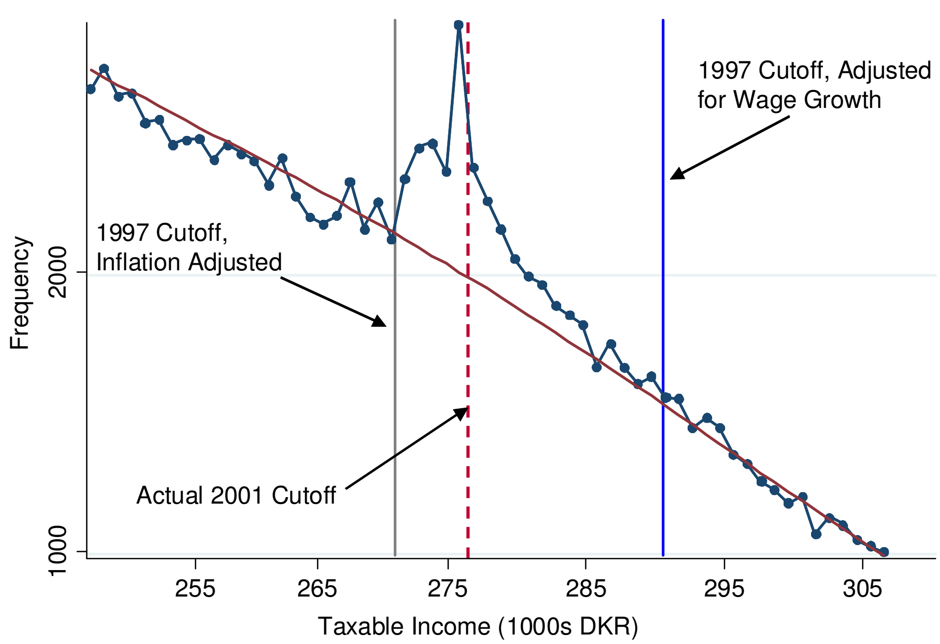
FIGURE 7

Distinguishing Changes in the Kink from Inflation and Wage Growth

**(a) Married Women, 1997**



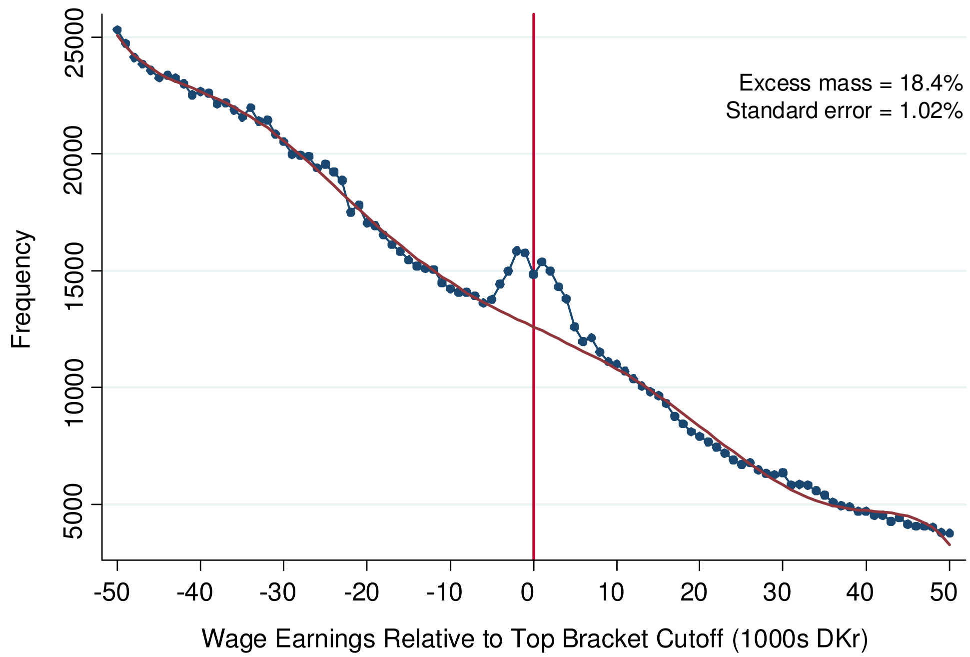

**(b) Married Women, 2001**



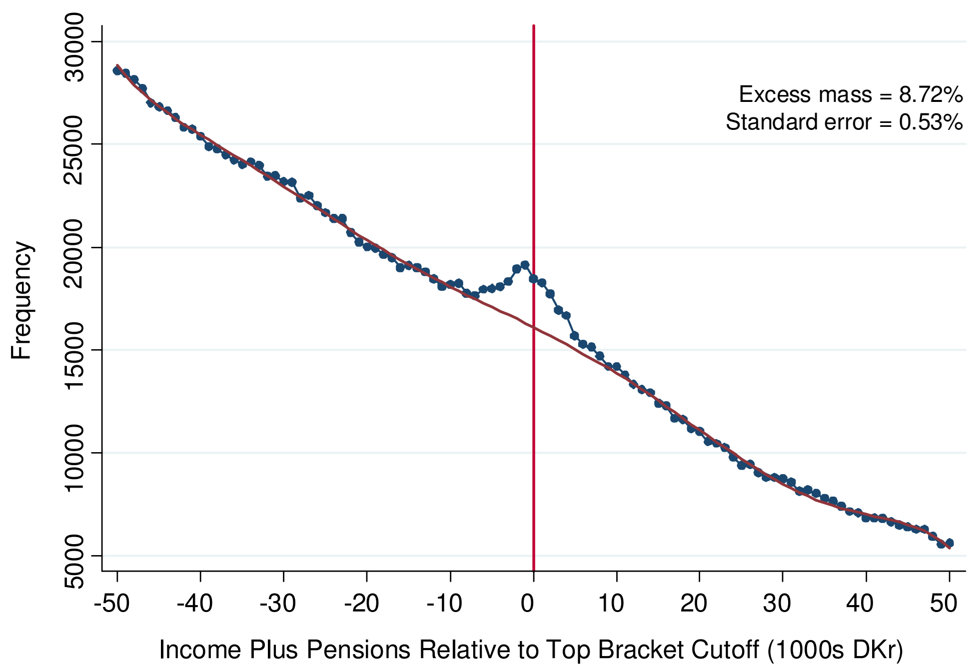

Notes: Panel (a) replicates the income distribution in Figure 6d, zooming in around the top tax bracket. The location of the bracket cutoff in 1997 is marked with the dashed line. The solid grey line shows the level of the 1994 top bracket cutoff adjusted for inflation; the solid blue line shows the 1994 bracket adjusted for average wage growth (right-most solid line). Panel (b) replicates (a) for the year 2001; the solid vertical lines reflect the 1997 bracket cutoff adjusted for inflation and average wage growth.

FIGURE 8
Tests for Income Shifting

**(a) Married Women: Wage Earnings**

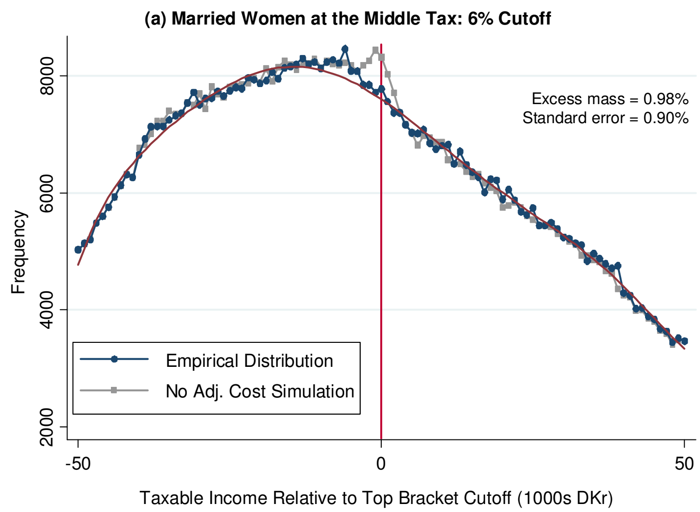


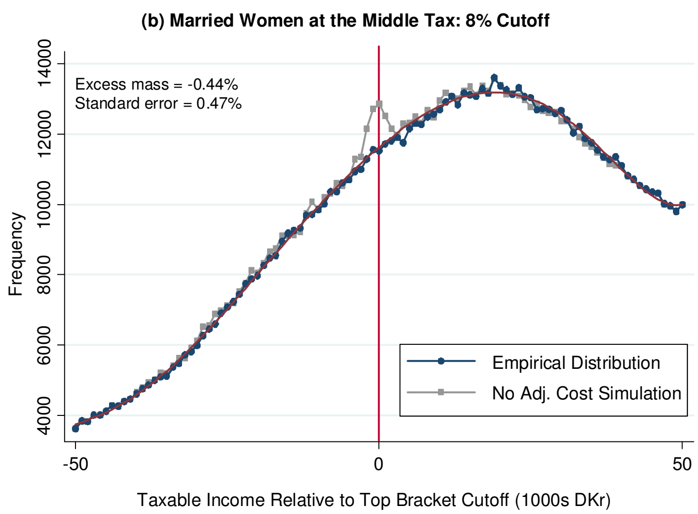

**(b) Married Women: Taxable Income plus Pensions**

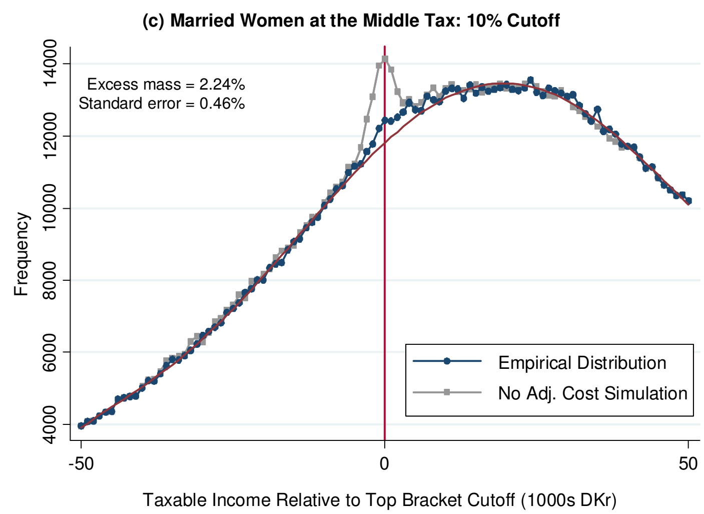



Notes: These figures plot the empirical distribution of (a) wage earnings and (b) taxable income plus contributions to both employer-run and private pension plans, as well as the counterfactual distributions and excess masses for wage earners who are married women between 1994-2001. For the methodology for calculating the counterfactual distributions, see the notes to Figure 4.

FIGURE 9
Bunching at Smaller Kinks

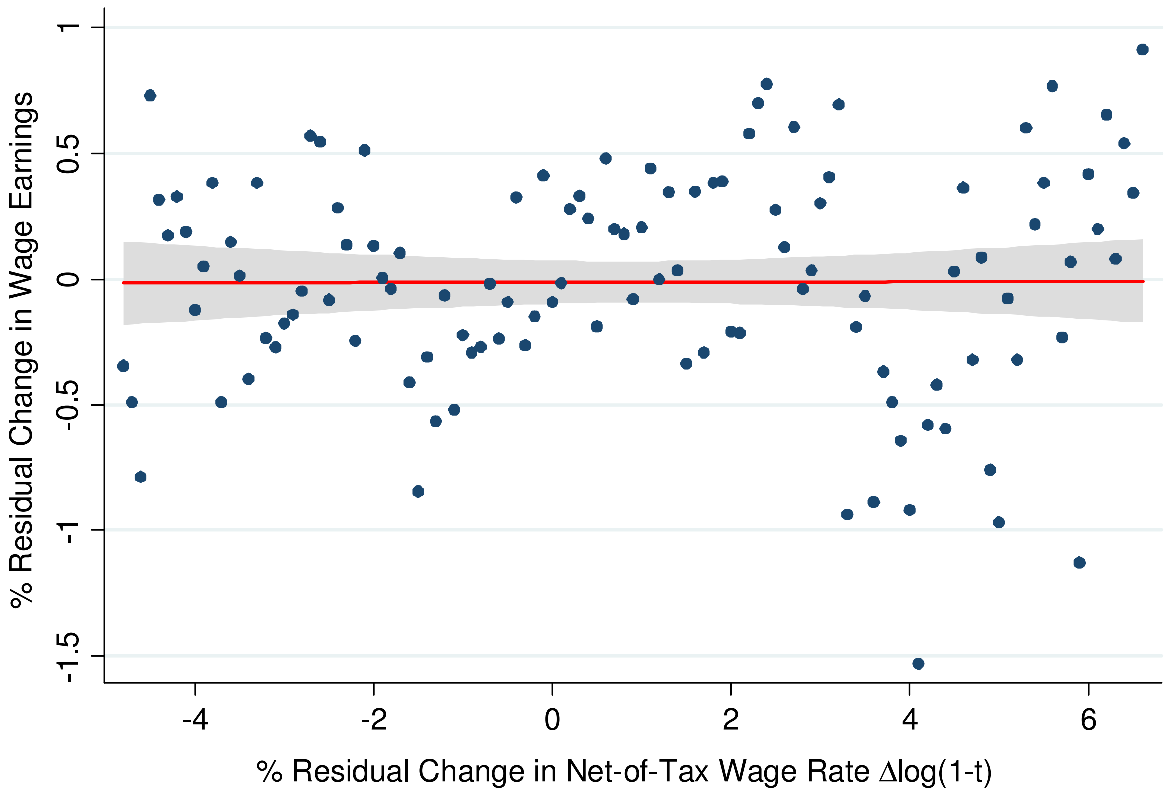


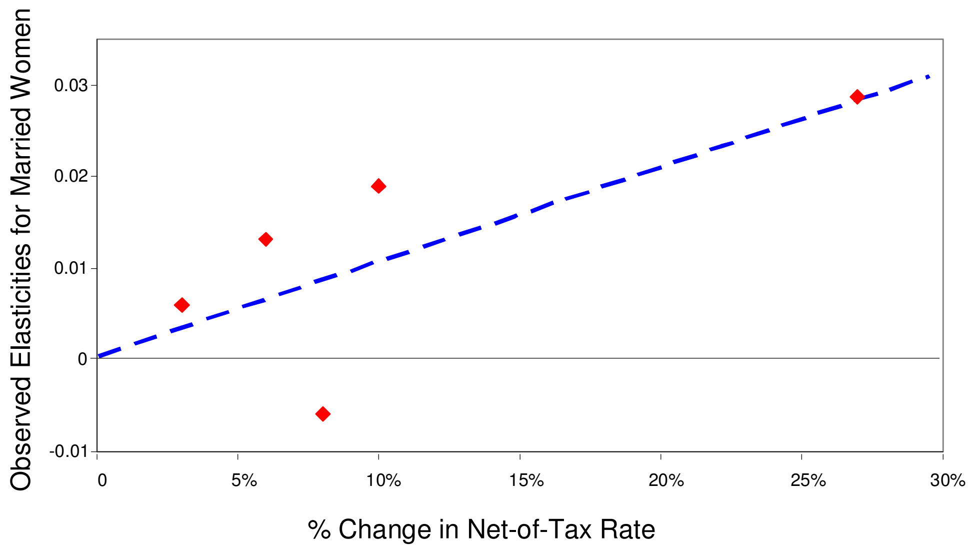

Notes: These three figures plot the empirical distribution of taxable income around (a) the lower middle tax bracket cutoff in 1994-1995, (b) the upper middle tax bracket cutoff in 1994-1995, and (c) the middle tax bracket cutoff in years 1996-2001, as well as the counterfactual distributions and excess masses, for married female wage earners. The figures additionally include a distribution showing the amount of bunching predicted if the elasticity were the same as that measured at the top bracket cutoff ($\varepsilon = 0.05$). For the methodology for calculating the counterfactual distributions, see the notes to Figure 4.

FIGURE 10
Observed Earnings Responses to Small Tax Reforms



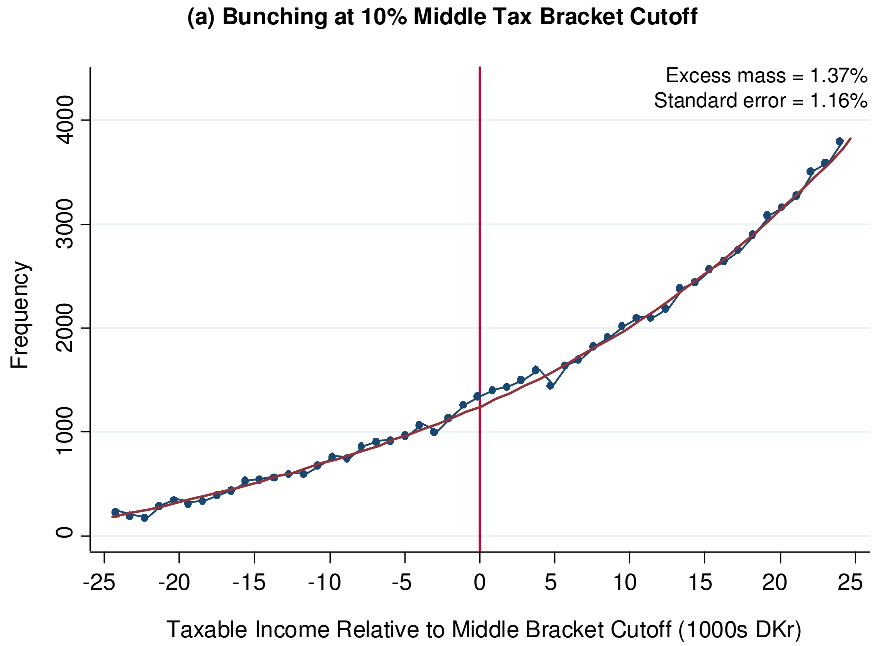

Notes: This figure plots the reduced-form relationship between simulated changes in net-of-tax wages and changes in wage earnings for female wage earners. We first regress $\Delta\log(1 - MTR_{it}^{\text{sim}})$ on a 10-piece wage earnings spline, a 10-piece spline in total personal income, and year, age, occupation, and region fixed effects. We then calculate residuals from this regression. Similarly, we regress $\Delta\log(y_{it})$ on the same set of variables and compute residuals. Finally, we bin the residual changes in simulated net-of-tax rates into 0.1% intervals. The figure plots the mean residual change in wage earnings in each bin. The red solid line shows the estimated relationship between the residuals, and the grey shaded region shows the 95% confidence interval for this fitted line.

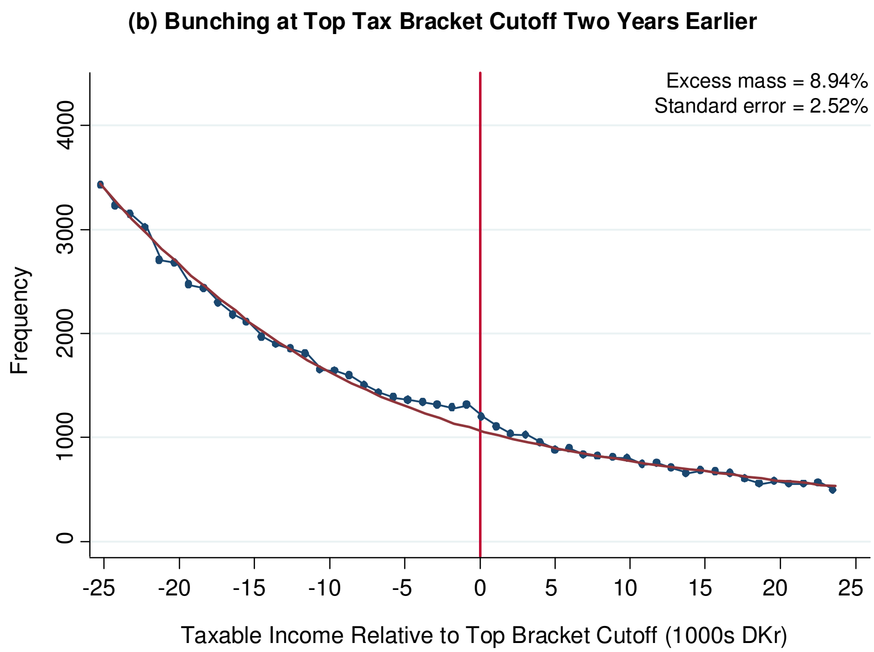
FIGURE 11
Observed Elasticities vs. Size of Tax Changes




Notes: This figure plots observed elasticities for married women against the percent change in the net-of-tax rate from the tax reforms, the three middle tax bracket cutoffs, and the top tax bracket cutoff. These data points correspond to the elasticity estimate in Column 3 of Table 2 and the excess masses in Figures 9a-9c and Figure 5a, respectively.

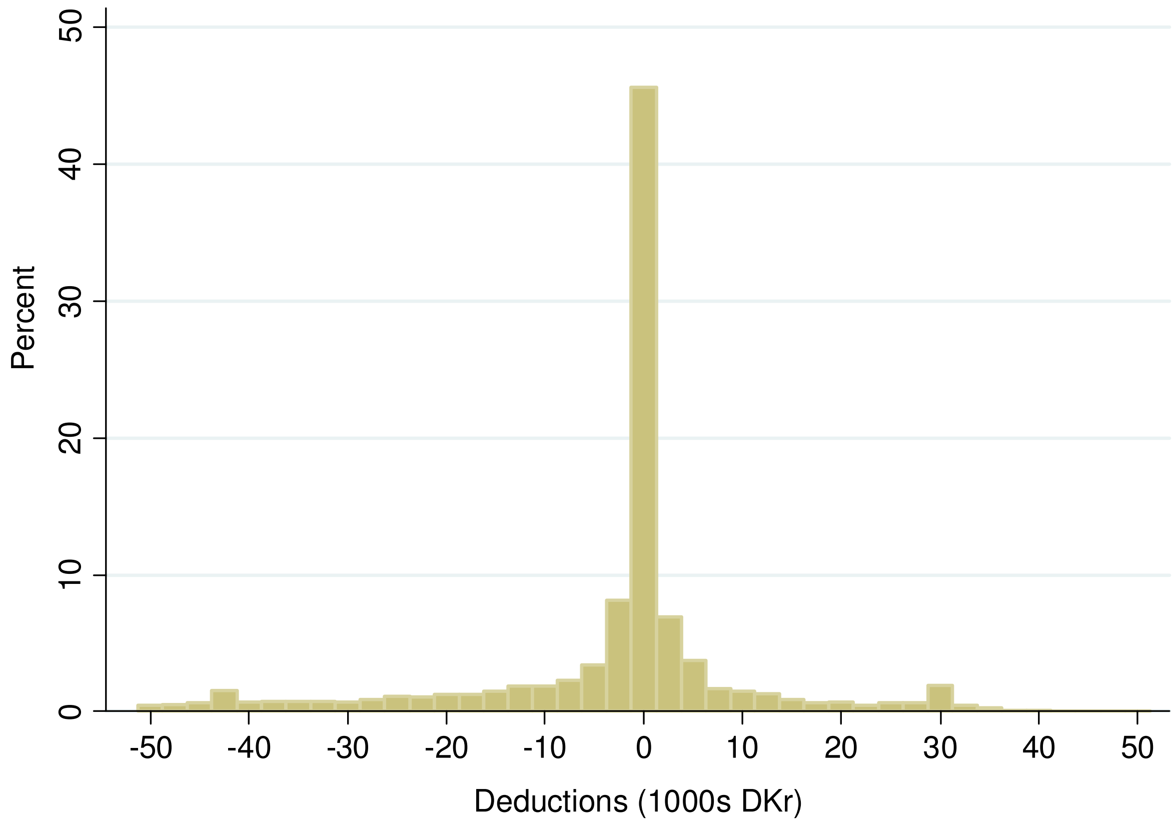
FIGURE 12
Bunching for Individuals who Switch Between Top and Middle Kinks




Notes: In these two figures, we identify a group of married women who earned within DKr 25,000 of the top tax bracket cutoff in one year and within DKr 25,000 of the middle tax bracket cutoff two years later. For this group, these two figures plot the empirical distribution of (a) taxable income in the earlier year and (b) income taxable above the middle tax bracket cutoff in the later year, as well as the counterfactual distributions and excess masses, for wage earners who are married women in each year from 1994-2001. For the methodology for calculating the counterfactual distributions, see the notes to Figure 4.

FIGURE 13
Distribution of Adjustments in 1995

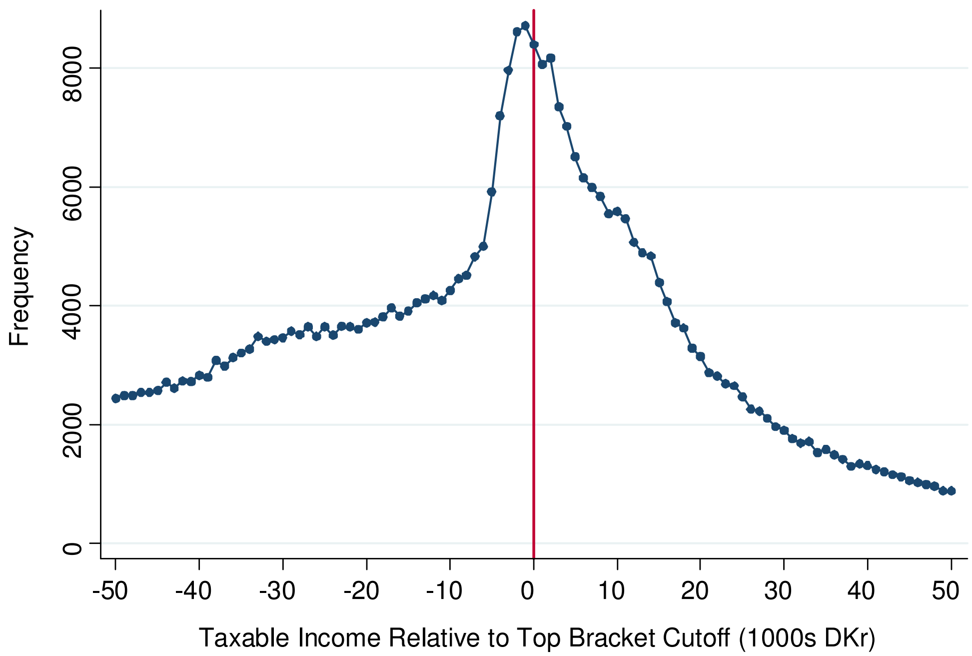



Notes: This figure plots a histogram of individual's deductions minus non-wage income in the top tax base in 1995.

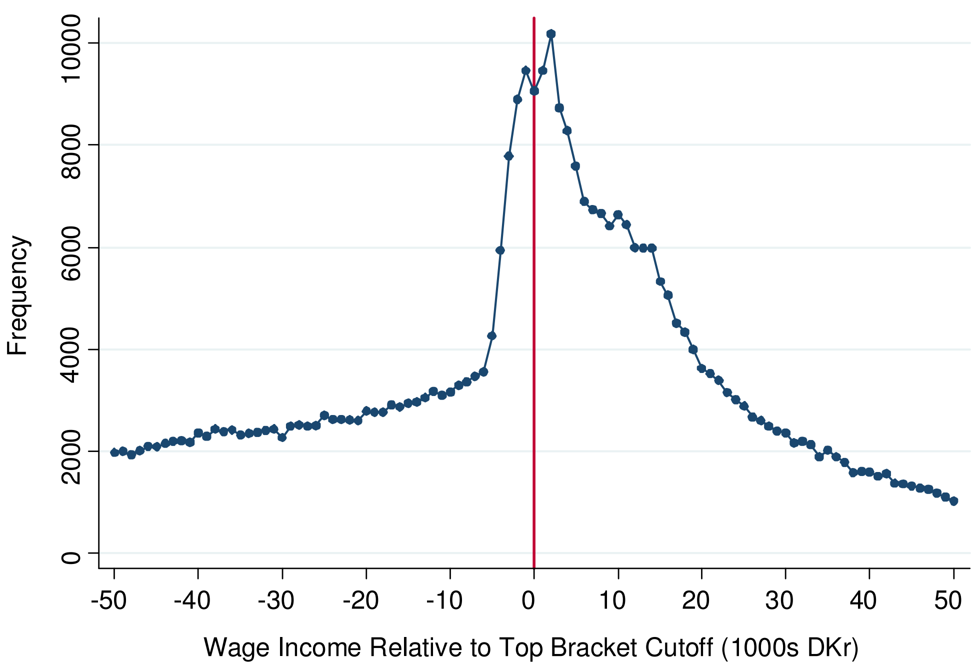
FIGURE 14

Teachers' Taxable Income and Wage Income Distributions

**(a) Teachers Taxable Income Distribution**

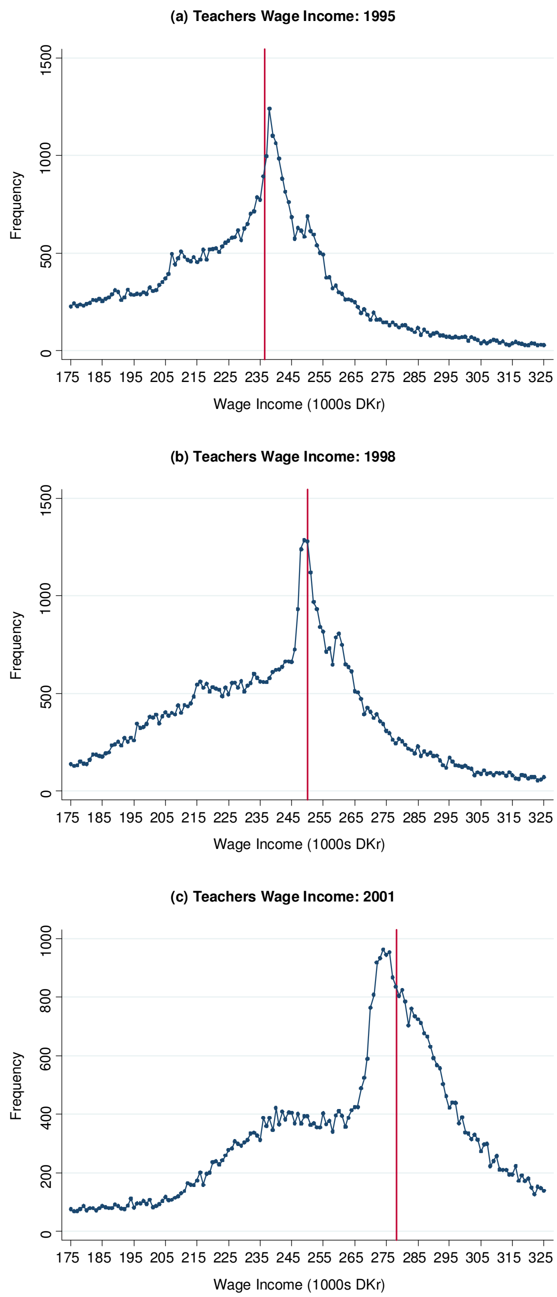



**(b) Teachers Wage Income Distribution**




Notes: These two figures plot the empirical distribution of (a) taxable income and (b) wage earnings relative to the statutory level of the top bracket cutoff for elementary school teachers (Danish ISCO 2331).

# FIGURE 15
Teachers Wage Income Distributions in 1995, 1998, and 2001

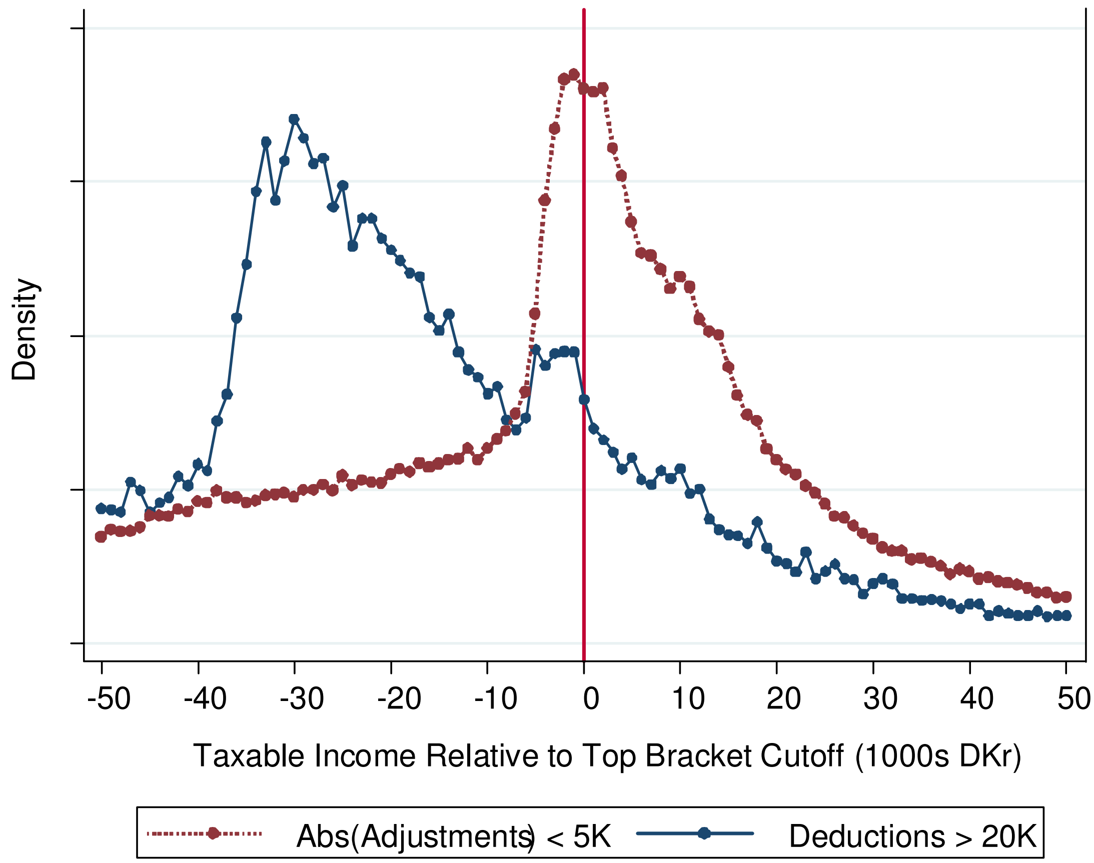



Notes: These three figures plot the empirical distribution of nominal wage earnings for elementary school teachers (Danish ISCO 2331) in (a) 1995, (b) 1998, and (c) 2001. The vertical lines in each figure mark the statutory level of the top tax bracket cutoff.

FIGURE 16
Wage Earnings Distribution: Teachers with Adjustments < -20,000

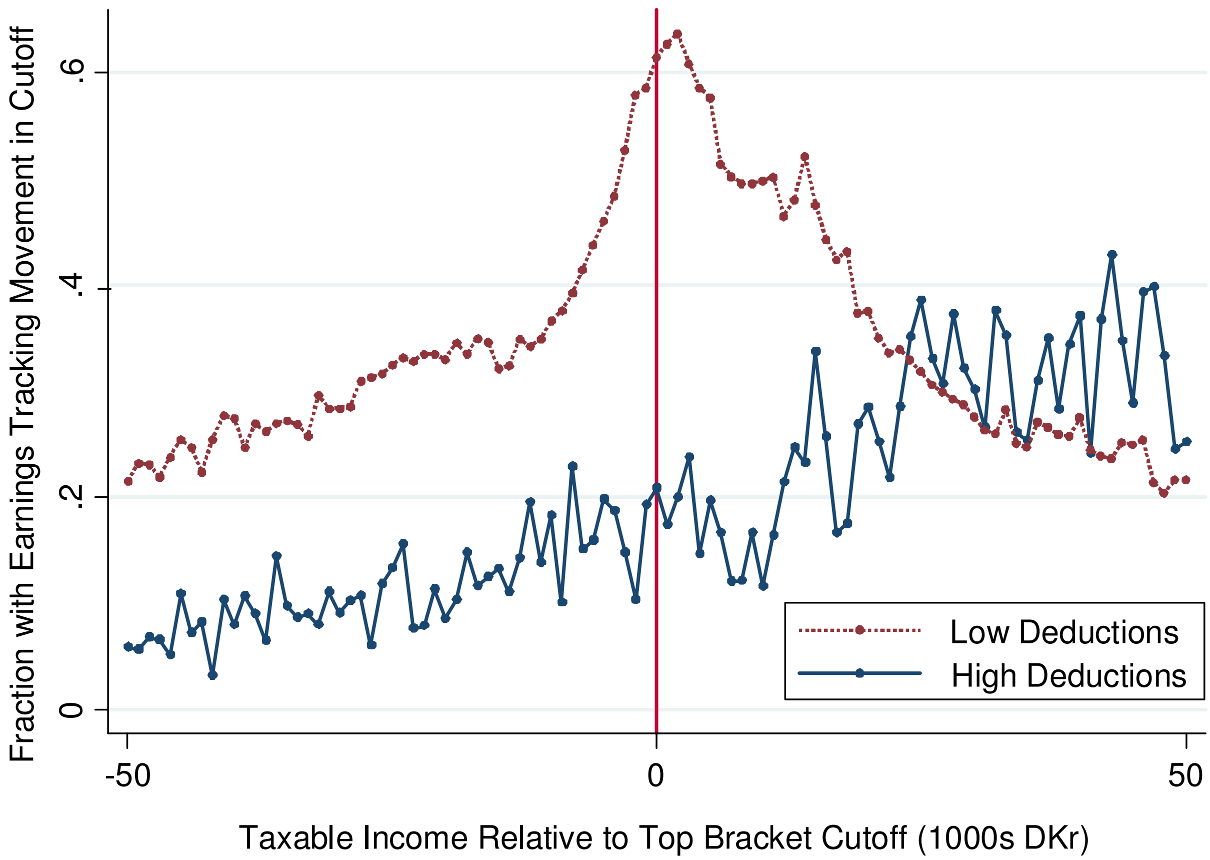

Notes: This figure plots the empirical distribution of wage earnings relative to the statutory level of the top tax bracket cutoff for wage earners between 1994 and 2001who are elementary school teachers, that is Danish ISCO 2331, and have non-wage earnings minus deductions equal to or less than DKr -20,000.

FIGURE 17
Teachers' Taxable Income Distributions: Effect of Adjustments



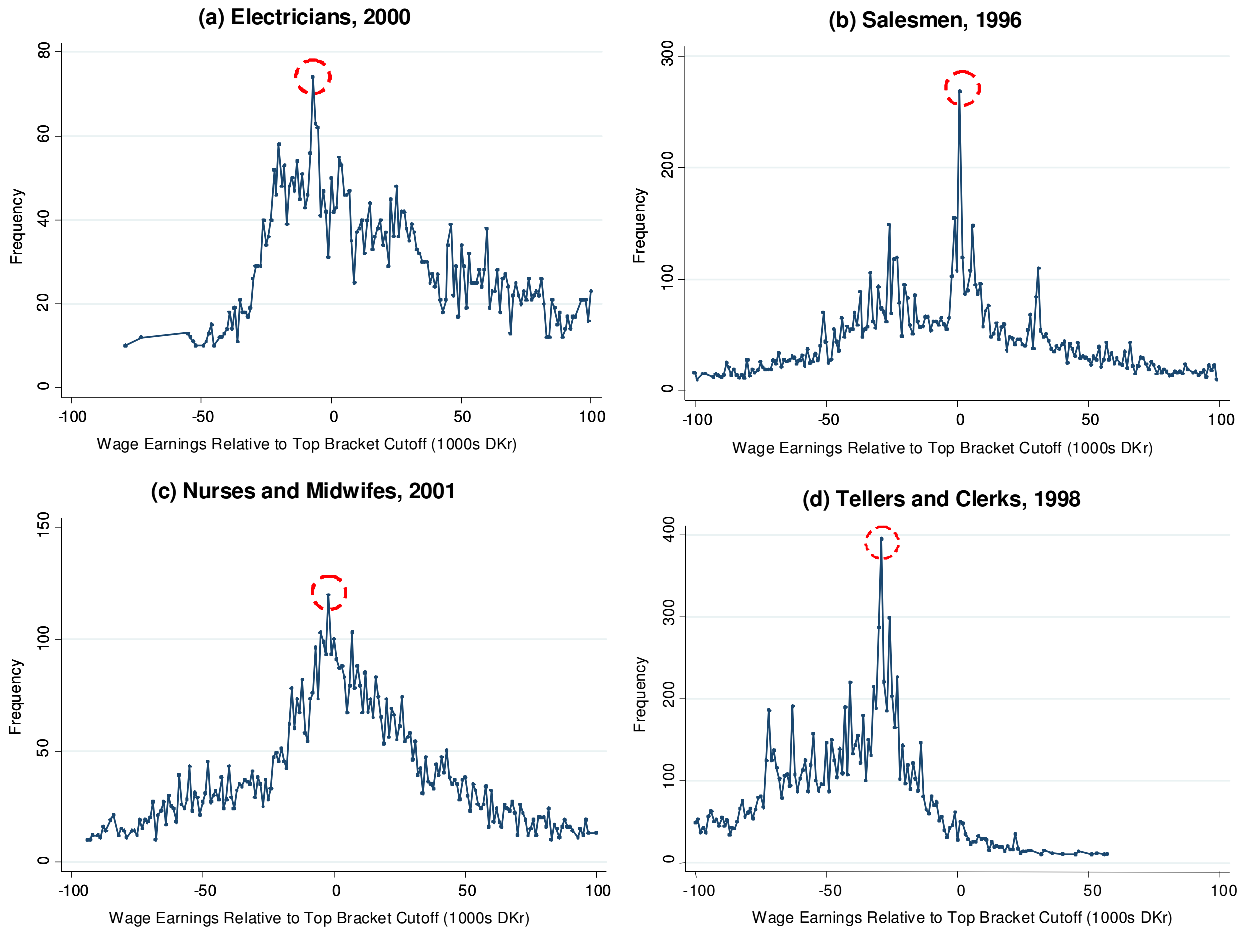

Notes: This figure plots the empirical distribution of taxable income relative to the statutory level of the top tax bracket cutoff for wage earners who are elementary school teachers, that is Danish ISCO 2331, within 1994-2001, for (dashed) those with total adjustments with DKr 5,000 of zero and (solid) those with total adjustments less than DKr -20,000.

FIGURE 18
Dynamics of Teachers' Earnings Around the Top Tax Bracket



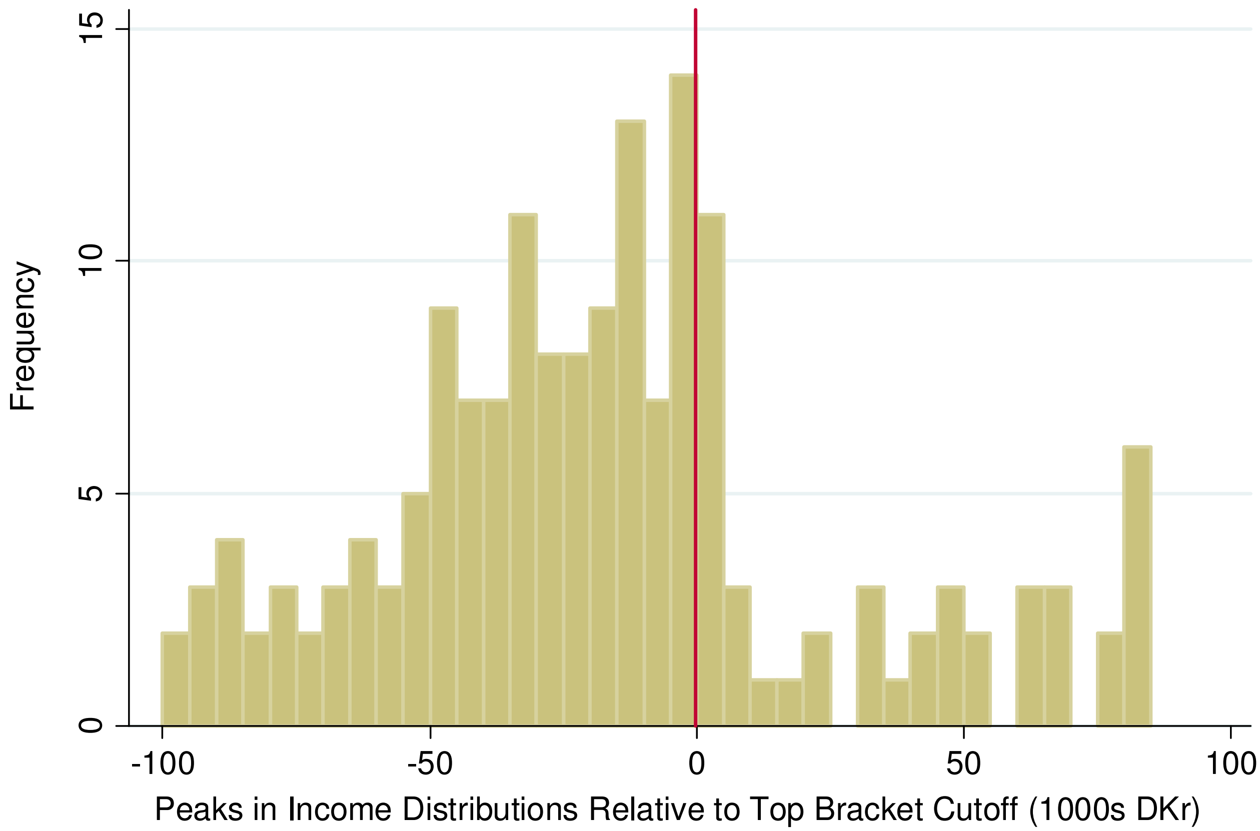

Notes: For each level of income relative to the top tax bracket cutoff in some year 0, on the *x*-axis, this figure plots the probability that wage earnings growth over two years is within DKr 7,000 of the two-year growth of the statutory level of the top tax bracket cutoff for wage earners who are elementary school teachers, that is Danish ISCO 2331, and have non-wage earnings minus deductions equal to or less than DKr -20,000 within 1994-2001.

FIGURE 19
Wage Earnings Distributions for Selected Occupations and Years




Notes: These four figures plot the empirical distributions of taxable income within DKr 1,000 bins relative to the top tax bracket cutoff for (a) electricians (Danish ISCO 3114) in 2000, (b) salesmen (ISCO 3419) in 1996, (c) nurses and midwives (ISCO 2230) in 2001, and (d) tellers and clerks (ISCO 4212) in 1998. The mode of each distribution is circled.

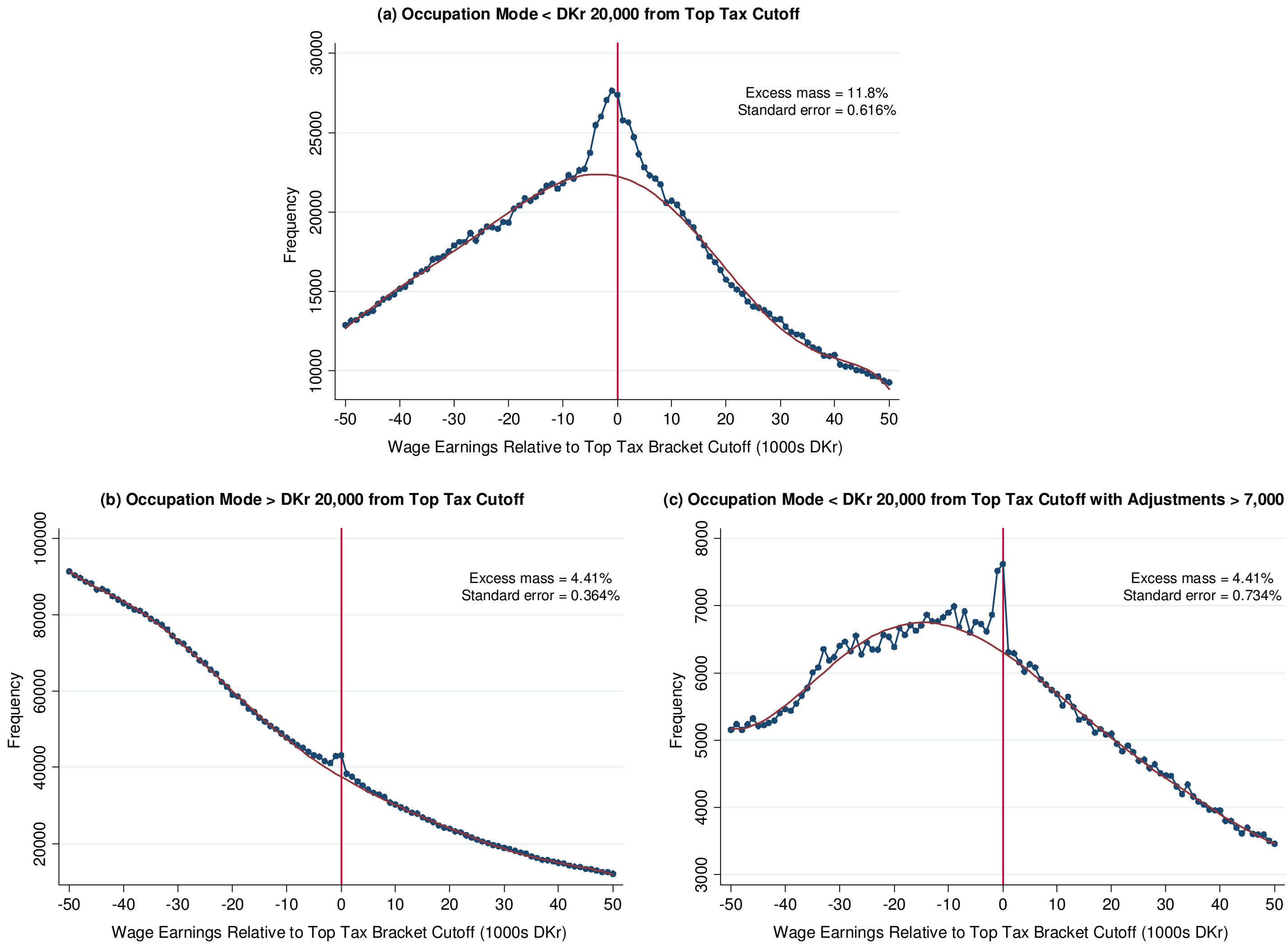
FIGURE 20
Distribution of Modes in Occupation Wage Earnings Distributions




Notes: To construct this figure, we first take each Danish ISCO occupation in each year that has more than 5,000 wage earners, and calculate the mode of each distribution (as in Figure 19). This figure plots a histogram of these peaks relative to the statutory level of the top tax bracket cutoff (with DKr 5,000 bins).

FIGURE 21

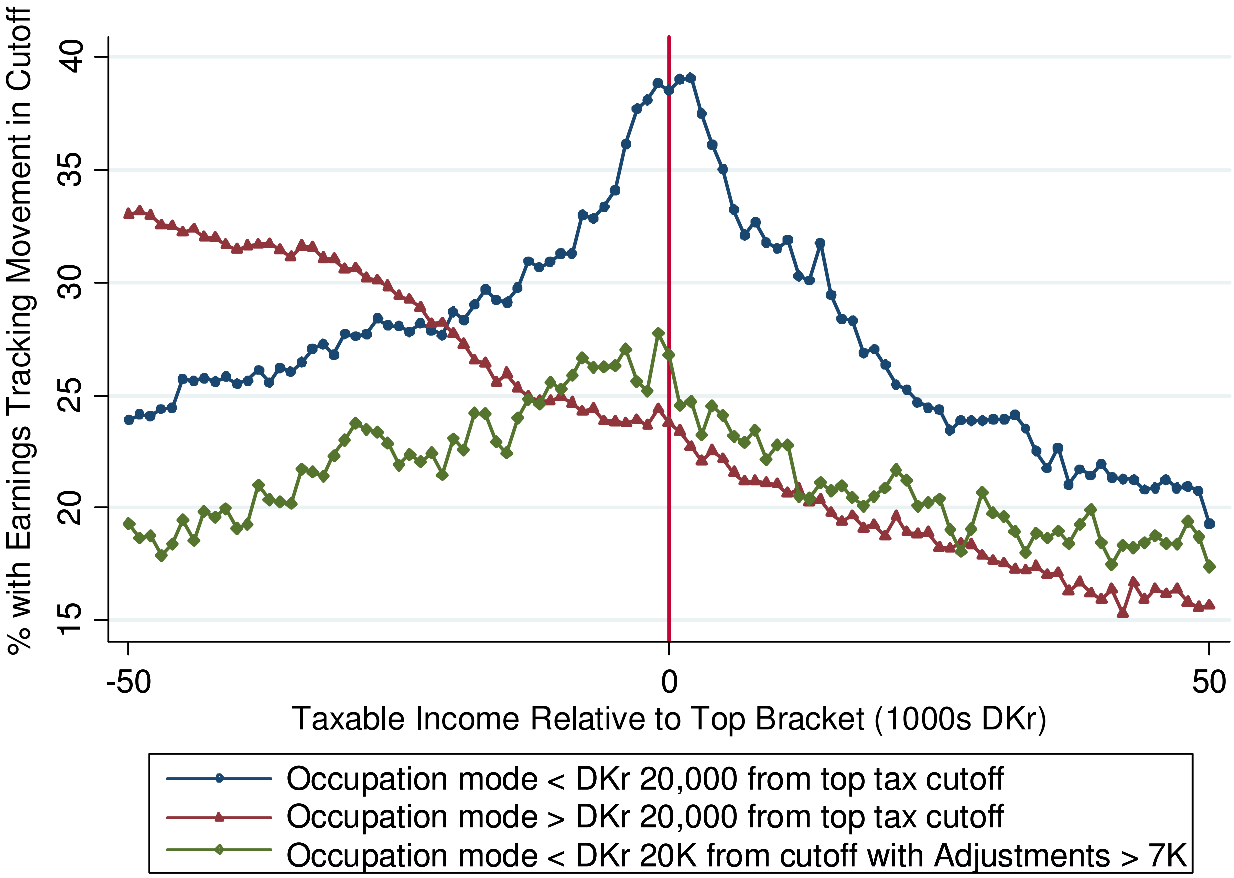
Effect of Occupation Modes on Wage Earnings Distributions around Top Tax Cutoff




Notes: In these three figures, we plot the empirical distribution of wage earnings for those workers in occupations (Danish ISCO codes) in which (a) the occupation mode is within DKr 20,000 of the statutory level of the top tax bracket cutoff, (b) the occupation mode is more than DKr 20,000 from the statutory level of the top tax bracket cutoff and (c) those in Figure (a) who also have total adjustments less than DKr 7,000 in absolute value. The figures also show the counterfactual distributions and excess masses. The sample is wage earners who are married women in each year from 1994-2001. For the methodology for calculating the counterfactual distributions, see the notes to Figure 4.

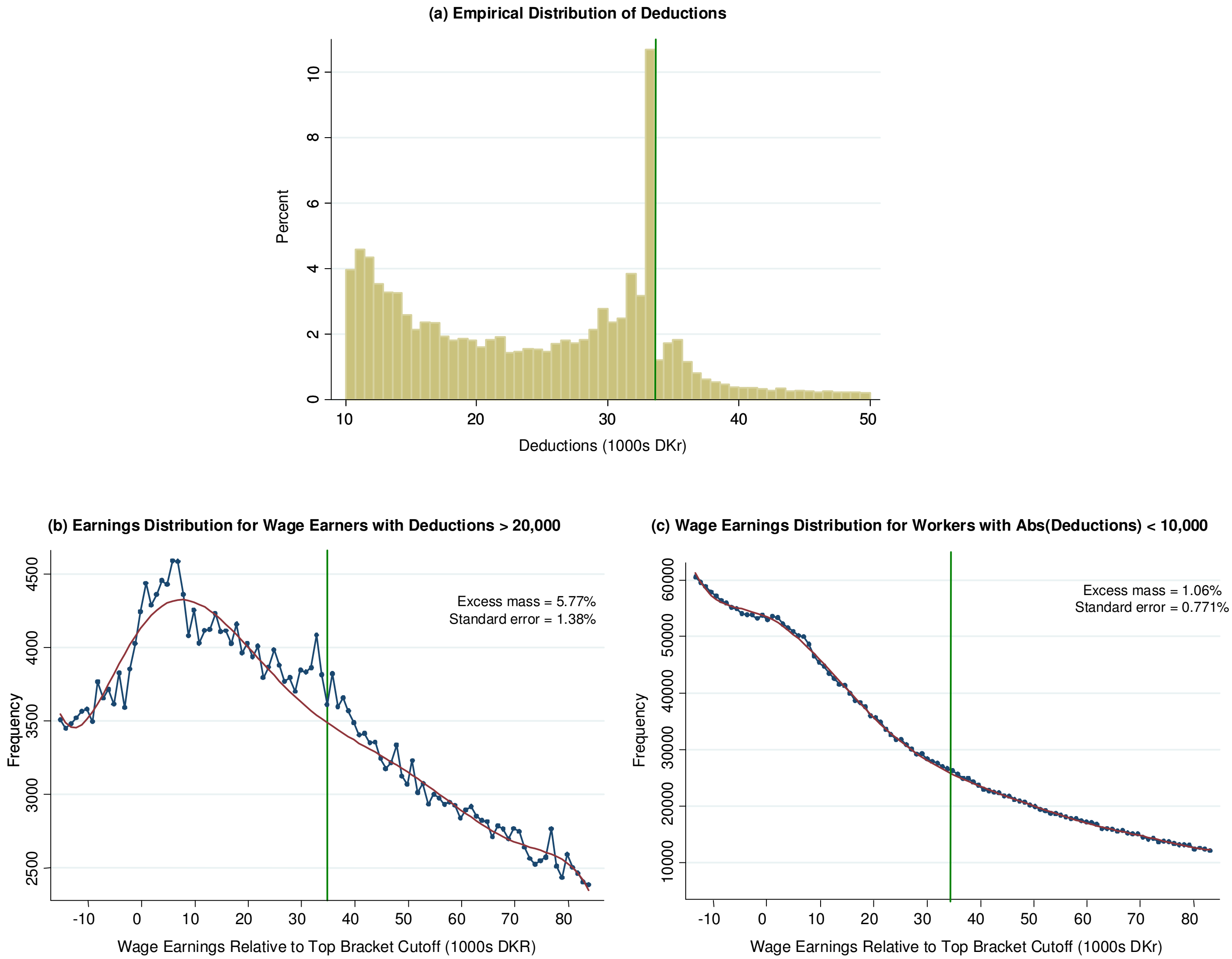
FIGURE 22

Dynamics of Earnings Around the Top Tax Bracket




Notes: This figure takes the three groups identified in Figures 21a-21c and plots, for each DKr 1,000 bin of taxable income in year 0, the probability that workers have two year wage earning growth within DKr 7,000 of the growth in the statutory level of the tax bracket cutoff over those two years.

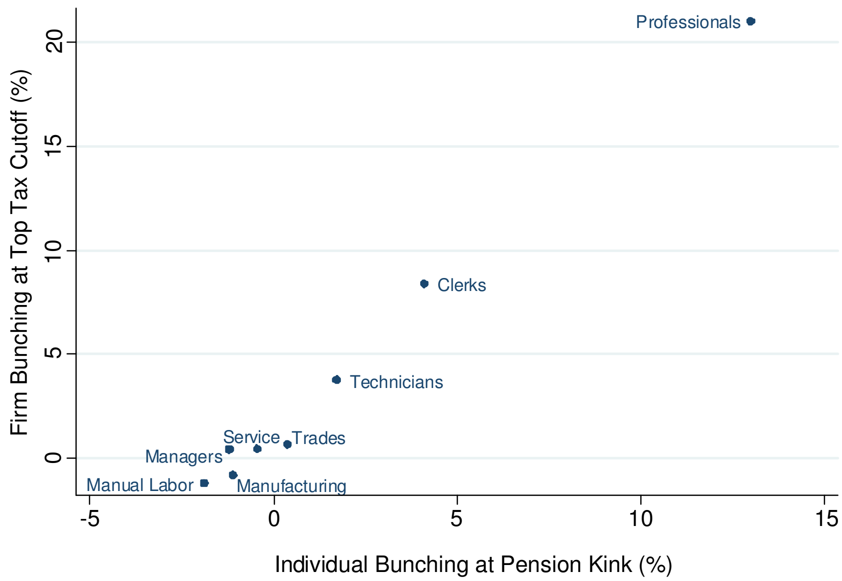
FIGURE 23
Individual Bunching in Wage Earnings



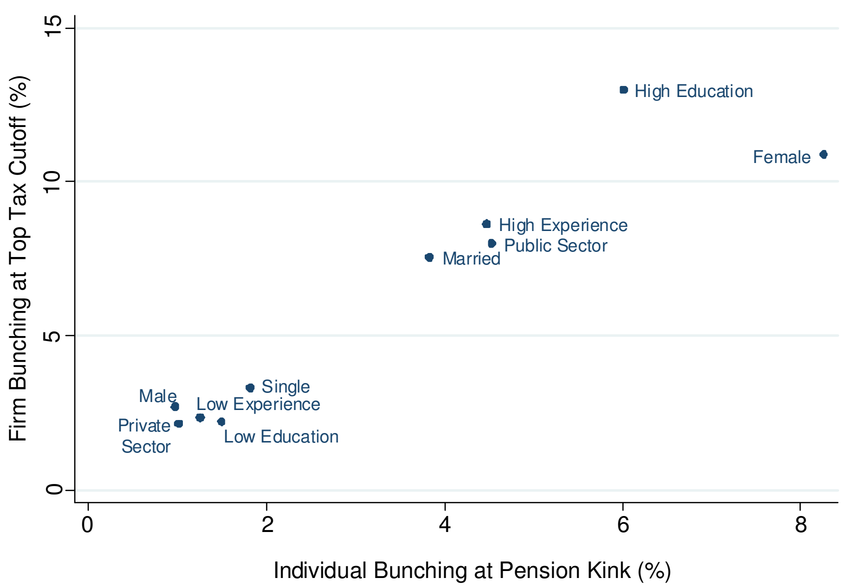

Notes: These figures plot (a) the distribution of adjustments to wage earnings for married women in 1994-2001, conditional on those adjustments being greater than DKr 10,000 and (b) the empirical distribution of wage earnings relative to the top tax bracket cutoff for married women in 1994-2001, along with the counterfactual distribution and excess mass. In Figure 23b, the counterfactual polynomial and excess mass apply not to the statutory level of the tax bracket cutoff. Instead, we estimate the excess mass around that level of wage earnings that would place a worker with deductions equal to the pension cap (an average of DKr 33,000) at the statutory level of the top tax bracket cutoff. In Figure 23c, we repeat the exercise in Figure 23b except for workers with less than DKr 10,000 in net deductions or non-wage income. For the methodology for calculating the counterfactual distributions, see the notes to Figure 4.

# FIGURE 24
## Correlation Between Individual and Firm Bunching

**(a) Individual and Firm Bunching by Occupation**



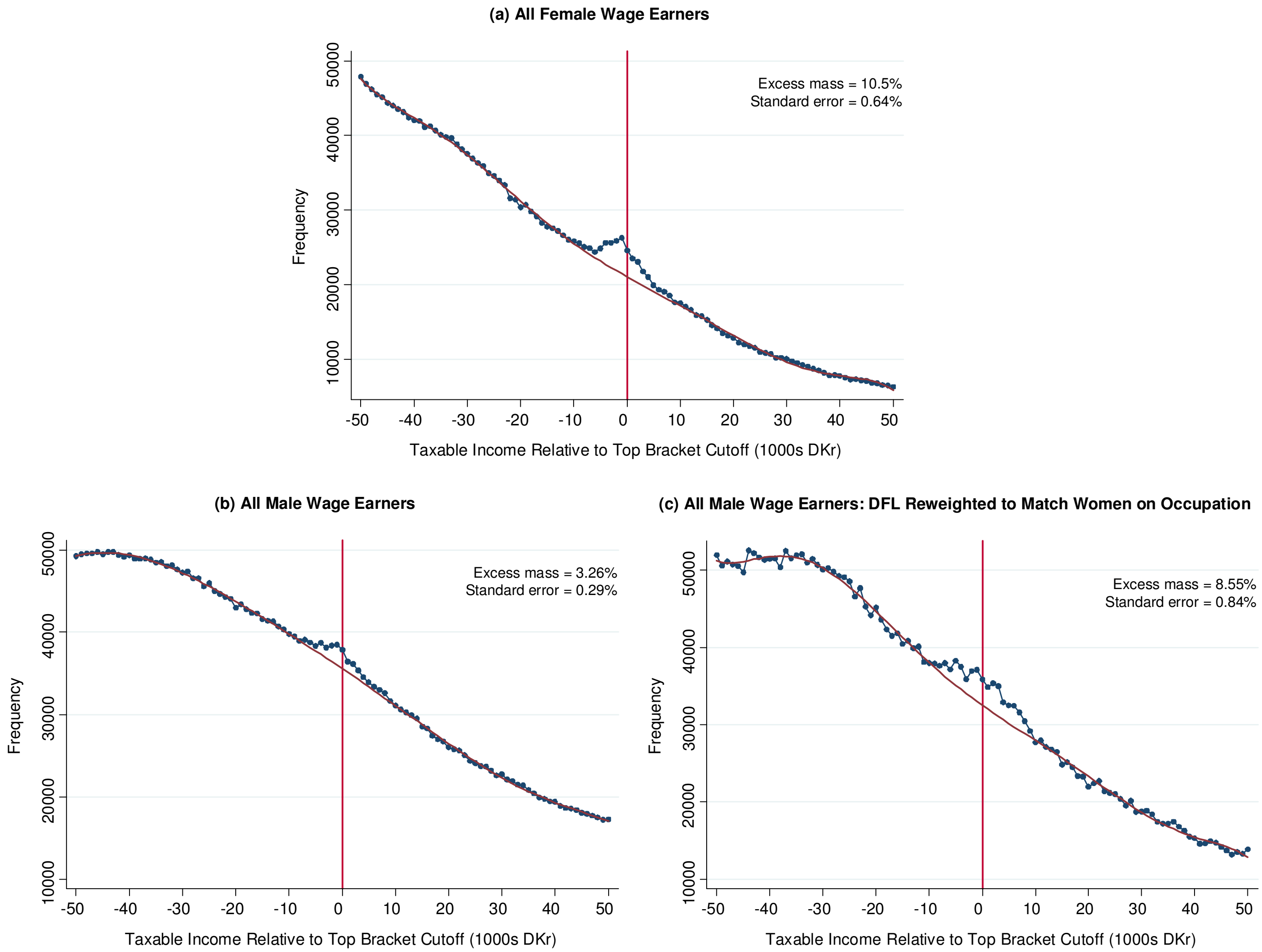

**(b) Individual and Firm Bunching by Demographic Groups**




Notes: These figures plot the amount of firm bunching vs. the amount of individual bunching for a variety of subgroups in the population. Individual bunching is the excess mass at the pension kink, as in Figure 23b. Firm bunching is excess mass at the top kink for those individuals with large deductions or non-wage income, as in Figure 16 (but here for all workers, not just teachers).

FIGURE 25

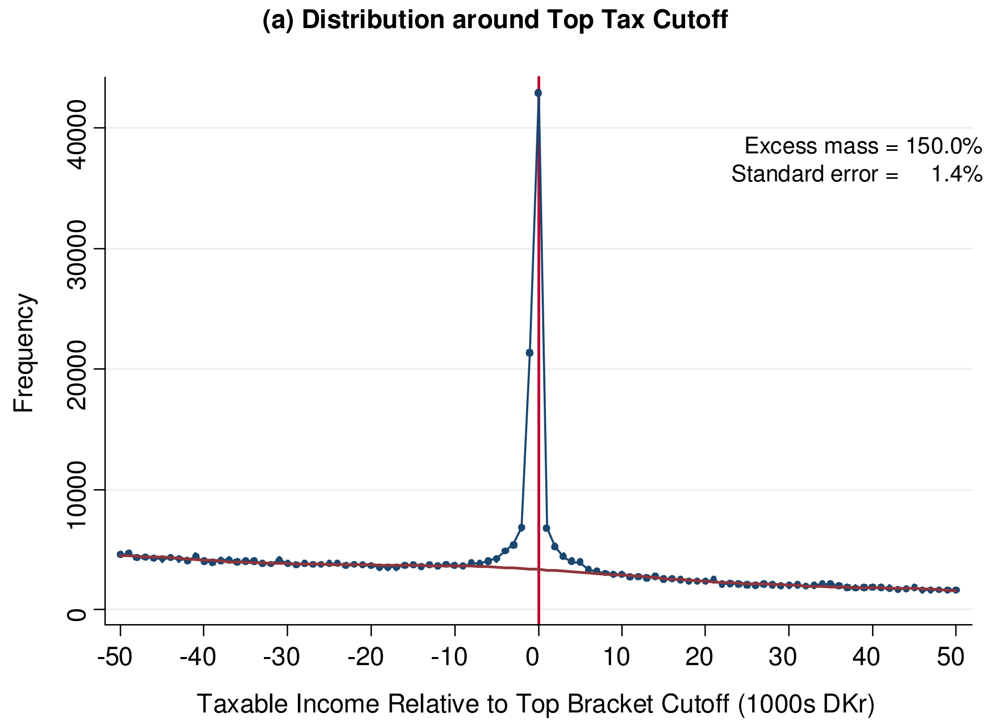
Male vs. Female Wage Earners: Effects of Occupational Heterogeneity

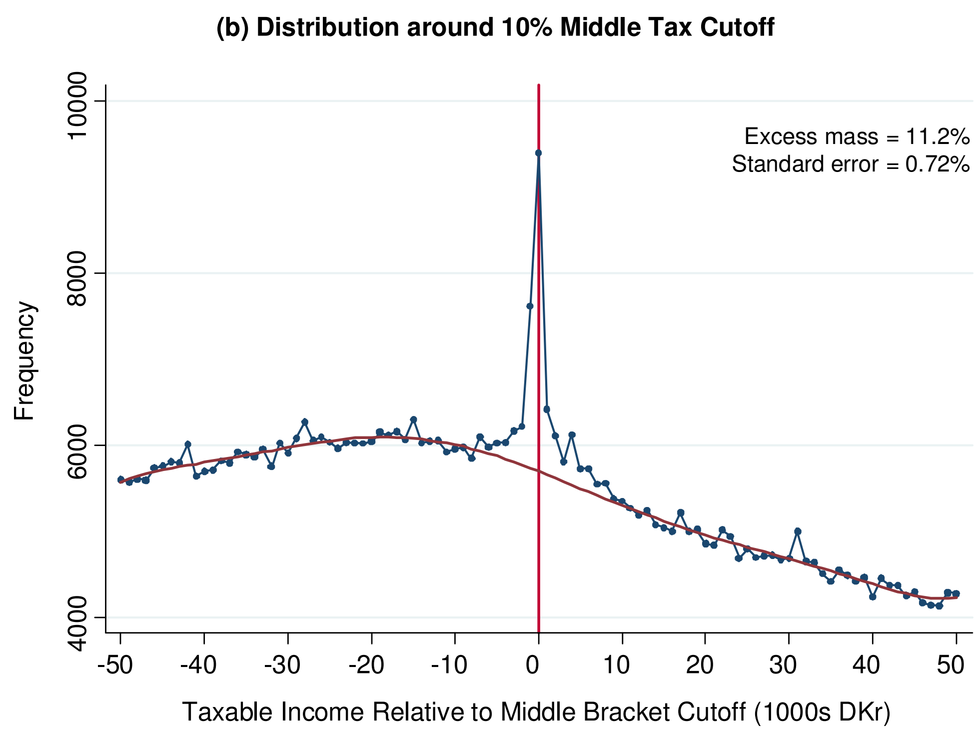



Notes: These figures plot the empirical distributions of taxable income in DKr 1,000 bins relative to the statutory level of the top tax bracket cutoff for 1994-2001 for (a) female wage earners and (b) male wage earners. Panel C replicates B, reweighting the observations to match the occupational distribution of those in Panel A. Following DiNardo, Fortin, and Lemieux (1996), we reweight an observation in occupation $i$ by $\frac{p_i}{1-p_i}$, where $p_i$ is the probability that a wage earner in occupation $i$ is female.

FIGURE 26

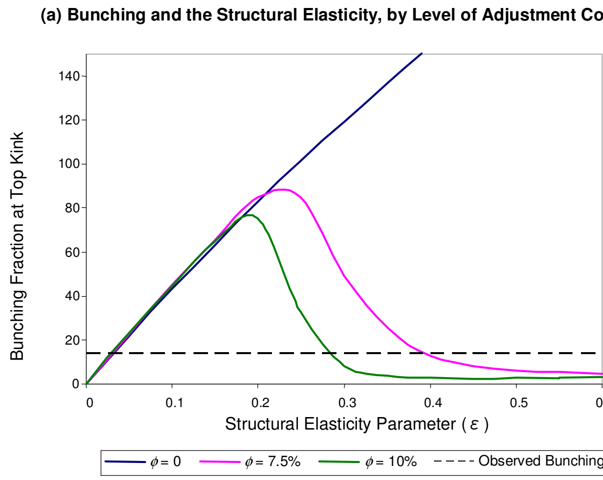
Taxable Income Distributions for Individuals with Self-Employment Income

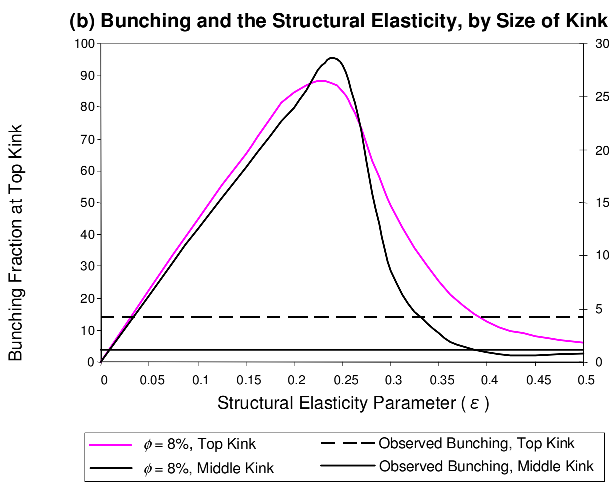


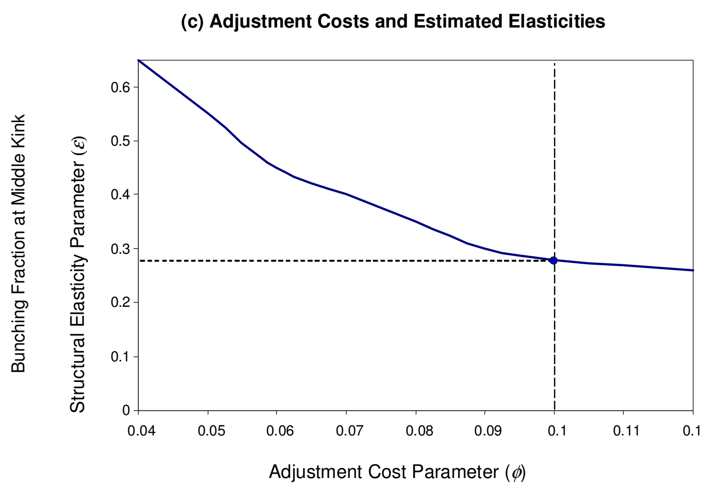




Notes: This figure plots the empirical distribution of taxable income from 1994-2001 around the top tax cutoff (Panel A) and 10% middle tax cutoff (Panel B) for individuals who report positive self-employment income on their tax returns.

# FIGURE 27
## Calibration of Model and Bound on Structural Elasticity




Notes: This figure shows how we calibrate the model to match the empirically observed amount of bunching at the top and middle kinks. Panel A plots the simulated relationship between the structural elasticity $\varepsilon$ and the amount of bunching at the top kink for three values of the adjustment cost: $\phi = 0\%$, $\phi = 8\%$, and $\phi = 10\%$. The horizontal dashed line in this figure shows the empirically observed amount of bunching at the top kink. Panel B plots the amount of bunching at the top and middle kinks vs. $\varepsilon$ for the $\phi = 8\%$ case. The horizontal lines show the amount of bunching at the two kinks; the figure illustrates that $\varepsilon = 0.4$ fits the data when $\phi = 8\%$. Panel C shows the value of $\varepsilon$ that fits the data for each value of $\phi$. The vertical line shows our upper bound on $\phi$ of 10%, which implies a lower bound on $\varepsilon$ of 0.29. See Appendix B for parametric assumptions and simulation methodology underlying these figures.

**Table I**

Summary Statistics for the Danish Population and Estimation Samples, 1994-2001

| | Population (1) | Wage Earners (2) | Self-employed (3) |
|---|---|---|---|
| Age | 41.01 | 38.16 | 47.39 |
| Number of children 0-17 years in household | 0.61 | 0.69 | 0.69 |
| Labor market experience (years) | 14.84 | 17.10 | 9.93 |
| White collar indicator | 23.25% | 36.39% | 0.00% |
| Post-compulsory Schooling | 60.95% | 68.63% | 63.84% |
| Higher education | 18.36% | 22.50% | 16.57% |
| Married | 47.87% | 49.00% | 67.97% |
| Female | 49.43% | 47.28% | 29.44% |
| | | | |
| Taxable Income | 185,749 | 221,227 | 184,740 |
| Wage Earnings | 144,526 | 211,037 | 21,379 |
| Adjustments | 41,224 | 10,191 | 163,361 |
| Capital Income | -12,916 | -17,346 | -8,503 |
| | | | |
| Pays the top tax (2000) | 20.50% | 27.06% | 25.67% |
| Pays the middle tax (2000) | 52.23% | 68.89% | 44.53% |
| | | | |
| Number of observations (1994-2001) | 30,176,237 | 19,383,681 | 1,721,415 |

NOTE--Table entries are means unless otherwise noted. Column 1 is based on the full population of Denmark from 1994-2001. Column 2 includes all wage earners (who report zero self-employment income on their tax returns), which is the primary estimation sample. Column 3 includes only individuals who report positive self-employment income. All monetary values are in real 2000 Danish Kroner. In the terminology of the Danish tax register, what we term "taxable income" is the top tax base equal to the sum of wage earnings and net positive capital income above a floor minus deductions; what we label "wage earnings" is labor income net of the gross tax. Finally, we use "adjustments" to refer to the sum of non-labor, non-capital income (e.g. self-employment income) minus deductions such as pension contributions.

**TABLE 2**
**Observed Elasticity Estimates using Small Tax Reforms**

| | *Dependent Variable: Percent Change in Labor Income* | | | | |
|---|---|---|---|---|---|
| Subgroup: | All Female Wage Earners | | Married Females | Married Fem. Professionals w/ High Exp. | Female Teachers |
| | (1) | (2) | (3) | (4) | (5) |
| % change in net-of-tax rate ($\Delta$ log (1-t)) | -0.014<br>(0.008) | 0.002<br>(0.006) | 0.006<br>(0.010) | 0.009<br>(0.033) | -0.024<br>(0.031) |
| Labor income spline | x | x | x | x | x |
| Total income spline | x | | x | x | x |
| Year fixed effects | x | x | x | x | x |
| Age fixed effects | x | x | x | x | x |
| Region fixed effects | x | | x | x | x |
| Occupation fixed effects | x | | x | x | x |
| Observations | 6,281,767 | 6,286,833 | 3,203,742 | 212,815 | 253,283 |

Notes: Standard errors clustered by id reported in parentheses. Dependent variable in all specifications is two-year growth rate in labor income. Independent variable is two-year growth rate in net-of-tax rate, instrumented using two-year growth rate in simulated net-of-tax rate based on base-year variables. Estimates can be interpreted as observed wage earnings elasticities from tax reforms.