ADJUSTMENT COSTS, FIRM RESPONSES, AND MICRO VS. MACRO LABOR SUPPLY ELASTICITIES: EVIDENCE FROM DANISH TAX RECORDS*

RAJ CHETTY
JOHN N. FRIEDMAN
TORE OLSEN
LUIGI PISTAFERRI

We show that the effects of taxes on labor supply are shaped by interactions between adjustment costs for workers and hours constraints set by firms. We develop a model in which firms post job offers characterized by an hours requirement and workers pay search costs to find jobs. We present evidence supporting three predictions of this model by analyzing bunching at kinks using Danish tax records. First, larger kinks generate larger taxable income elasticities. Second, kinks that apply to a larger group of workers generate larger elasticities. Third, the distribution of job offers is tailored to match workers' aggregate tax preferences in equilibrium. Our results suggest that macro elasticities may be substantially larger than the estimates obtained using standard microeconometric methods. *JEL* Codes: H20, J20.

I. INTRODUCTION

The vast theoretical and empirical literature on taxation and labor supply generally assumes that workers can freely choose jobs that suit their preferences. This paper shows that the effect of taxes on labor supply is shaped by two factors that limit workers' ability to make optimal choices: adjustment costs and hours constraints determined endogenously in equilibrium. We present quasi-experimental evidence showing that these forces attenuate microeconometric estimates of labor supply elasticities.

To motivate our empirical analysis, we develop a stylized labor supply model with job search costs and endogenous hours

*We would like to thank David Card, Stephen Coate, Edward Glaeser, James Hines, Han Hong, Lawrence Katz, Henrik Kleven, Claus Kreiner, Patrick Kline, Erzo Luttmer, Robert Moffitt, John Pencavel, Emmanuel Saez, Laszlo Sandor, Esben Schultz, anonymous referees, 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, Peter Elmer Lauritsen at Statistics Denmark, 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, Jane Choi, Jessica Laird, and Keli Liu provided outstanding research assistance. Support for this research was provided by the Robert Wood Johnson Foundation and NSF Grant SES-0645396.

The Quarterly Journal of Economics (2011) 126, 749–804. doi:10.1093/qje/qjr013.

constraints. We model hours constraints by assuming that each firm requires its employees to work a fixed number of hours because of an ex-ante commitment to a production technology. Workers draw offers from the aggregate distribution of hours and can search for jobs that offer hours closer to their unconstrained optimum by paying search costs. We consider two types of equilibrium in the labor market: competitive markets and collective bargaining. In the competitive case, both workers and firms are price takers. In the collective bargaining case – which is more relevant for our empirical application – unions bargain with firms over wages and the aggregate hours distribution. Under both notions of equilibrium, the number of jobs posted by firms at each level of hours must equal the number of workers who select those hours after the search process is complete. The aggregate distribution of workers' preferences therefore determines the hours constraints imposed by firms in equilibrium. However, most individuals do not work their unconstrained optimal number of hours because of search costs.

Our model produces a divergence between macro labor supply elasticities (defined as the effect on average hours of work of variation in taxes across economies) and micro labor supply elasticities (defined as the effect of tax changes or kinks in non-linear tax systems that affect subgroups of workers). We show that the macro elasticity always equals the "structural" labor supply elasticity ε, the parameter of individuals' utility functions that determines elasticities absent frictions. In contrast, micro elasticities are attenuated relative to ε because of search costs and hours constraints.

The model generates three testable predictions about how search costs and hours constraints affect the labor supply (or taxable income) elasticities observed in micro studies. All three predictions hold irrespective of whether the labor market equilibrium is determined by competition or collective bargaining. The first prediction is that the observed elasticity increases with the size of the tax variation from which the estimate is identified. Intuitively, large tax changes prompt more individuals to pay search costs and find a new job. Analogously, larger kinks induce more individuals to pay search costs to find a job that places them at the kink. Second, the observed elasticity increases with the number of workers affected by a tax change or kink. Changes in taxes induce changes in labor supply not just by making individuals search for different jobs, but also by changing the equilibrium distribution

of hours. Because changes in taxes that affect a larger group of individuals induce larger changes in hours constraints – either through market forces or directly through unions – they generate larger observed elasticities. Furthermore, tax changes may 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 individual responses to tax and responses to taxes induced by aggregation of workers' tax preferences through firms or unions. In particular, one should observe larger distortions in the equilibrium distribution of job offers in sectors or occupations where workers themselves exhibit larger tax elasticities.

We test these three predictions using a matched employer-employee panel of the population in Denmark between 1994 and 2001. This dataset combines administrative records on earnings and taxable income, demographic characteristics, and employment characteristics such as occupation and tenure. 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 10% or less, and changes in tax rates across tax brackets within a year, which generate variation in net-of-tax wages of up to 35%. We focus primarily on the cross-bracket variation in taxes rates because it is larger and applies to large subgroups of the population, permitting coordinated responses. In particular, we estimate taxable income elasticities by measuring the amount of bunching at kink points, as in Saez (2010).[1]

Consistent with the first prediction, the elasticities implied by the amount of bunching at large kinks are significantly larger than those implied by the amount of bunching at smaller kinks. There is substantial, visually evident excess mass in the wage earnings distribution around the cutoff for the top income tax bracket in Denmark, at which the net-of-tax wage rate falls by approximately 30%. There is little 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%.

1. Following the modern public finance literature reviewed in Saez, Slemrod, and Giertz (2009), we proxy for "labor supply" using taxable income. We discuss the implications of measuring taxable income elasticities instead of hours elasticities below.

Similarly, we find no changes in earnings around the small tax reforms that change net-of-tax wages by less than 10%. The observed elasticities at the largest kinks are several times larger than those generated by smaller kinks and tax reforms across a broad range of demographic groups, occupations, and years. Using a series of auxiliary tests, we show that the differences in observed elasticities are driven by differences in the size of the tax changes rather than heterogeneity in elasticities by income levels or tax rates.

To test the second prediction, we exploit heterogeneity in deductions across workers. In Denmark, 60% of wage earners have zero deductions. These workers reach the top tax bracket when their wage earnings exceeds the top tax cutoff for taxable income, which we term the "statutory" top tax cutoff. Workers with large 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 firms and unions cater to the tax incentives of the most common workers. In particular, the mode of occupation-level wage earnings distributions has an excess propensity to be located near the statutory top tax cutoff.[2] Importantly, the wage earnings distribution even for workers who have substantial deductions or non-wage income exhibits excess mass at the *statutory* top tax cutoff. Because these workers do not face any change in marginal tax rates at the statutory cutoff, this finding constitutes direct evidence that wage-hours offers are tailored to the tax preferences of the majority of workers who have small deductions. We label this supply-side response to tax incentives induced by the aggregation of workers' tax preferences "aggregate bunching".

Although aggregate bunching is an important source of behavioral responses to the tax system, some of the bunching 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 3% of workers make pension contributions up to this amount and therefore cross into the highest income tax

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.

bracket when they earn DKr 33,000 more than the statutory top tax cutoff. We find 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 appears to be driven solely by individual job search, as there is no excess mass at the pension-driven kink for workers with small deductions. Because of aggregate bunching, workers with common tax preferences (those with small deductions) have a higher propensity to bunch at the top tax kink than those with uncommon tax preferences (those with large deductions).

We test the third prediction by estimating the correlation between individual and aggregate bunching across occupations. We find that there is more bunching at the statutory kink in occupations where workers exhibit more individual bunching in wage earnings at the pension-driven kink. Although this result cannot be interpreted as a causal effect because the variation in individual bunching is not exogenous, it is consistent with the prediction that firms and unions cater to workers' aggregate tax preferences.

All of the results above are obtained for wage earners. We analyze self-employed individuals separately. As the self-employed do not face significant adjustment costs or hours constraints, one would expect that none of our three predictions should hold for this subgroup. Indeed, we find that the self-employed exhibit sharp bunching at both small and large kinks, show no evidence of aggregate bunching at the statutory kink, and are equally likely to bunch irrespective of their deductions. These placebo tests support our hypothesis that search costs and hours constraints are the key factors that attenuate micro elasticity estimates for wage earners.

Although our findings show that adjustment costs and hours constraints are likely to dampen observed elasticities, they do not identify the underlying structural elasticity ε relevant for macro comparisons. Identifying ε would require estimating a structural model of labor supply with frictions and endogenous hours constraints. Such an analysis is outside the scope of this paper, but two observations suggest that the structural elasticity ε is likely to be an order of magnitude larger than the observed elasticities in our data, which are below 0.02. First, calibrations of our stylized model consistently imply values of ε an order of magnitude larger than the observed elasticities at the top kink (Chetty

et al. 2009). Second, the self employed exhibit much larger taxable income elasticities than wage earners, suggesting that individuals do seek to optimize relative to taxes when they face fewer frictions.[3]

Our results could help explain why macro studies find much larger elasticities than microeconometric studies (Blundell and MaCurdy 1999; Saez, Slemrod, and Giertz 2009; Chetty 2011).[4] Micro estimates are attenuated by frictions because they are identified from individuals' responses to changes in tax rates or kinks *after* obtaining a job near their optimum. In contrast, macro variation in tax rates across countries changes the jobs individuals search for and the jobs offered by firms to begin with, producing larger elasticities.[5] Our explanation for the gap between micro and macro elasticities complements recent work arguing that macro elasticities are larger because they incorporate both extensive and intensive margin responses (e.g. Rogerson and Wallenius 2009). Much of the difference in labor supply across countries with different tax regimes is driven by hours worked conditional on employment (Davis and Henrekson 2005; Chetty et al. 2011). That is, macro estimates of *intensive margin* elasticities are much larger than their microeconometric counterparts. Our analysis explains this divergence between intensive margin elasticities. We caution, however, that our findings do not provide justification for the very large elasticities (e.g. $\varepsilon > 1$) used in some macro models.

In addition to the literature on micro vs. macro elasticities, our study builds on and contributes to several other strands of the literature on labor supply. First, previous work has proposed that adjustment costs and hours constraints affect labor supply decisions (e.g. Cogan 1981; Ham 1982; Altonji and Paxson 1988; Dickens and Lundberg 1993; Rogerson 2005) and that long-run

3. This finding is consistent with a recent literature that documents larger elasticities for workers who can control their hours more easily, such as stadium vendors (Oettinger 1999), bike messengers (Fehr and Goette 2007), and cab drivers (Farber 2005).

4. A recent microeconometric study that uses the same Danish microdata as we do here (Kleven and Schultz 2010) estimates an elasticity of zero by studying tax reforms over a twenty year period.

5. Frictions could also explain why macro studies find large (Frisch) elasticities when analyzing fluctuations in labor supply over the business cycle. Intertemporal wage fluctuations are large for certain subgroups and much of the fluctuation in hours at business cycle frequencies is on the extensive rather than intensive margin (Chetty 2011).

elasticities may differ from short-run elasticities (Holmlund and Söderström 2008).[6] Our contribution is to show how these factors affect estimates of intensive-margin labor supply elasticities using quasi-experimental methods. Our findings also support the hypothesis that the effects of government policies may operate through coordinated changes in social norms or institutions rather than individual behavior (e.g. Lindbeck 1995; Alesina, Glaeser, and Sacerdote 2005).

Second, our results contribute to the literature on non-linear budget sets (e.g., Hausman 1981; Moffitt 1990; MaCurdy, Green, and Paarsch 1990), where the lack of bunching at kinks creates problems in fitting models to the data. 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.

Third, our analysis relates to recent work on taxable income as a measure of labor supply (Feldstein 1999; Slemrod and Yitzhaki 2002; Chetty 2009). The bunching we observe is driven by changes in wage earnings rather than tax avoidance via pension contributions or evasion. However, because our dataset does not contain information on hours of work, we cannot rule out the possibility that some of the responses we observe arise from income shifting. Importantly, distinguishing income shifting from hours of work is not critical for the conclusions we draw here, as our three predictions also apply to an environment with adjustment costs and coordination constraints in income shifting.

The paper is organized as follows. In Section II, we set up the model, define micro and macro elasticities formally, and derive the testable predictions. Section III describes the Danish data and provides institutional background. Section IV presents the empirical results. Section V concludes.

6. Our paper differs from the recent work of Chetty (2011) in two ways. First, while Chetty (2011) derives bounds on elasticities under the assumption that individuals face adjustment costs, we provide direct empirical evidence that adjustment costs affect observed elasticities within a single economy. Second, Chetty (2011) focuses exclusively on worker behavior, while we model endogenous hours constraints and firm/union responses in equilibrium.

II. Search Costs and Hours Constraints in a Labor Supply Model

This section develops a stylized model of labor supply on the intensive-margin whose purpose is to highlight the channels through which frictions affect labor supply elasticities. We analyze a static model because our empirical analysis focuses on how search costs and hours constraints interact in equilibrium rather than on the dynamics of adjustment in labor supply. We present some results on responses to tax reforms in a two-period extension of this stylized model in Online Appendix A.[7]

II.A. Setup

Firms. Firms have one-factor linear production technologies. Each firm employs a single worker to produce goods sold at a fixed price p. Let $w(h)$ denote the hourly wage rate paid to workers who work h hours in equilibrium. Firm j posts a job that requires h_j hours of work at the wage rate $w(h_j)$. We model hours constraints by assuming that a firm cannot change the hours it posts after matching with a worker.[8] This assumption captures the intuition that firms sink capital in a technology that requires a certain amount of labor for production *before* hiring workers. Such constraints may emerge from technological benefits of coordinating work schedules (as in an assembly line), the fixed costs of restructuring job and benefit packages, or regulations such as overtime pay requirements.[9]

A firm that posts a job with h_j hours earns profit

$$\pi_j = ph_j - w(h_j)h_j.$$

Let the aggregate distribution of hours offered by firms be denoted by a cdf $G(h)$. A key feature of our model is that the aggregate

7. All appendix material is available online at http://qje.oxfordjournals.org/.

8. This model is isomorphic to one in which a single firm offers heterogeneous hours packages and workers face costs of switching jobs within the firm. This is because the boundary of a firm is indeterminate with constant returns to scale.

9. We focus on hours constraints in the model for simplicity, but they should be interpreted more broadly as technological constraints on job characteristics (e.g. training, effort, benefit packages).

distribution of hours constraints $G(h)$ is *endogenously* determined in equilibrium, as we describe below.[10]

Workers. Workers, indexed by i, have quasi-linear utility

$$u_i(c,h) = c - \alpha_i^{-1/\varepsilon}\frac{h^{1+1/\varepsilon}}{1+1/\varepsilon} \tag{1}$$

over a numeraire consumption good c and hours of work h. The heterogeneous taste parameter $\alpha_i > 0$, is distributed according to a smooth cdf $F(\alpha_i)$ with full support on a closed interval. This utility specification eliminates income effects and generates a constant wage elasticity of labor supply ε in a frictionless model. We abstract from income effects because the variation in marginal tax rates at kinks that we exploit for identification has little effect on average tax rates and thus generates negligible income effects. We extend the analysis to utility functions that generate non-constant elasticities in Online Appendix A.

To characterize tax changes that affect subgroups of the population differently, assume that there are two types of tax systems, indexed by $s \in \{NL, L\}$.[11] Individuals with $s_i = NL$ face a two-bracket non-linear tax system with marginal tax rates of τ_1 and $\tau_2 > \tau_1$. These workers begin to pay the higher tax rate when their incomes $w_i h_i$ exceed a threshold K. Individuals with $s_i = L$ pay a linear tax rate of τ on all income. With this tax system, individual i has consumption

$$c_i(h_i) = \begin{cases} (1-\tau_1)\min(w_i h_i, K) + \\ \quad (1-\tau_2)\max(w_i h_i - K, 0) & \text{if } s_i = NL \\ (1-\tau)\, w_i h_i & \text{if } s_i = L \end{cases} \tag{2}$$

A fraction ζ of workers face the non-linear tax system NL and the remainder $(1-\zeta)$ face the linear tax system L. Let worker i's optimal level of hours be denoted by $h_i^* = \arg\max_{h_i} u_i(c(h_i), h_i)$. The tax systems workers face are uncorrelated with their tastes: $F(\alpha_i | s_i) = F(\alpha_i)$.

10. This endogenous determination of wage-hours offers differentiates this model from the few existing models of hours constraints, in which firms' technologies exogenously determine the distribution of wage-hours packages (e.g. Rosen 1976).

11. For example, tax systems often treat single and married individuals differently, in which case the two types in our model would be defined by marital status.

Workers begin their search for a job by drawing an initial offer h_i^0 from the aggregate offer distribution $G(h)$. Each worker can either accept this offer or turn it down and search for another job. We assume that workers who search locate their optimal job h_i^*, but must pay a utility cost of search ϕ_i. As a result, workers will search for their optimal job if and only if the gains from the switch are larger than ϕ_i. This job search process for workers can be viewed as a functional $\mathcal{F}$ that maps an aggregate distribution of hours posted by firms $G(h)$ and wage schedule $w(h)$ to a new distribution $\mathcal{F}(G(h), w(h))$.

II.B. Equilibrium

To demonstrate that our testable predictions apply to both competitive and unionized labor markets such as that of Denmark, we analyze two different equilibrium concepts – one based on collective bargaining and another based on market competition.

Model 1: Collective Bargaining. There is a single union that represents all the workers in the economy. As in Earle and Pencavel (1990), we assume that the union bargains with firms over both wages and hours, holding fixed the number of available jobs. The union's objective is to maximize its members' aggregate utility subject to the constraint that all members must find jobs (full employment). Since there are many firms and one union, the union makes a take-it-or-leave-it offer to all firms, who may accept or decline it individually. The workers then search for jobs as described above. If there are more workers than firms at a given hours level after the search process, jobs are randomly rationed to workers, and hence some workers are unemployed.

In equilibrium, unions determine the wage and the distribution of hours, subject to the constraints that firms must participate in the labor market and all workers are employed. Because labor demand is infinitely elastic, firms will not accept $w > p$, and the unions impose $w = p$. In order to satisfy the full employment constraint, the union must choose a distribution of jobs $G(h)$ satisfying the fixed-point condition $G^*(h) = \mathcal{F}(G^*(h), p)$. This condition ensures that the distribution of hours endogenously reflects the aggregate distribution of worker preferences. If many workers prefer to work 40 hours per week, the union bargains to induce many firms to offer jobs that require 40 hours of labor per week in equilibrium.

Model 2: Market Equilibrium. In a decentralized competitive equilibrium, firms post an hours offer h_j chosen to maximize profit:

$$\pi_j = ph_j - w(h_j)h_j. \tag{3}$$

Intuitively, firms seek to produce at an hours level where the supply of labor exceeds demand, allowing them to earn profits by paying a wage $w(h_j) < p$. Because firms are free to enter the market at any level of hours h_j, profits are bid to zero, implying that $w(h_j) = w = p$ for all h_j in equilibrium. Market clearing requires that the distribution of jobs initially posted by firms coincides with the jobs selected by workers at the wage rate $w = p$ after the job search process is complete, i.e. $G^*(h) = \mathcal{F}(G^*(h), p)$.

Both the market equilibrium and collective bargaining models generate a fixed wage $w = p$ and a distribution of hours $G^*(h)$ that endogenously reflects the preferences of workers while ensuring full employment. The only difference between the two models is the mechanism through which worker preferences are aggregated to generate $G(h)$: through firms in the market equilibrium model and through unions in the collective bargaining model. Because the two models generate the same equilibrium hours distribution, the predictions derived below apply to both institutional structures of the labor market. The two models of wage setting produce the same equilibrium because our model assumes that labor demand is infinitely elastic. However, the key mechanisms that drive our testable predictions would also operate in a more realistic setting in which the labor demand elasticity is finite and unions extract rents. In particular, unions would continue to aggregate the tax preferences of the workers they represent, leading to larger responses to tax changes that have large size and scope.

Our model should be viewed as representing the equilibrium in a given sector or occupation. It is straightforward to generate heterogeneous wage rates by introducing multiple sectors. Suppose there are Q different skill types of workers and Q types of corresponding output goods sold at prices $p_1, \ldots, p_Q$. Workers of type q can only work at firms that produce good q, so there is no interaction across the Q segments of the labor market. Within each sector one union bargains with firms to set an equilibrium wage rate $w_q = p_q$ and an equilibrium hours distribution determined by its workers' preferences according to the model above.

The following sections characterize the properties of the equilibrium hours distribution $G(h)$, focusing on the relationship between tax rates and labor supply. For analytical convenience, we derive the key predictions in a series of special cases.

II.C. Special Case 1: Benchmark Frictionless Model

In the frictionless model ($\phi_i = 0$), the structural preference parameter ε fully determines the effects of taxes on labor supply. This is because workers who face no search costs always choose their unconstrained optimal level of hours h_i^*. For workers with $s_i = L$, who face a linear tax τ, the optimal level of hours is $h_i^* = \alpha_i ((1-\tau)w)^{\varepsilon}$. The hours choices of workers who face the nonlinear tax system are given by

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

where $\underline{\alpha} = h_K / ((1-\tau_1)w)^{\varepsilon}$ and $\overline{\alpha} = h_K / ((1-\tau_2)w)^{\varepsilon}$. Workers with moderate disutilities of labor supply $\alpha_i \in [\underline{\alpha}, \overline{\alpha}]$ bunch at the kink because the net-of-tax wage falls at h_K.[12]

Now consider how variation in the linear tax rate τ affects labor supply. When subject to a higher tax rate, workers of type $s_i = L$ optimally reduce their work hours by

$$d \log h = \varepsilon \cdot d \log(1-\tau). \tag{5}$$

This equation shows that the elasticity of hours with respect to the net-of-tax rate $(1-\tau)$ coincides with the structural parameter ε in the frictionless model. We shall therefore refer to ε as the "structural" elasticity. Workers of type $s = NL$, who are unaffected by τ, do not change hours of work and can be used as a control group in an empirical study.

In our one-dimensional labor supply model, the hours elasticity coincides with the elasticity of taxable wage income (wh)

12. The logic for why a mass of workers bunch at the kink is captured by the following quote from a Danish construction worker interviewed by a member of the Danish Tax Reform Commission: "By the end of November, some of my colleagues stop working. It does not pay anymore because they have reached the high tax bracket."

with respect to the net-of-tax-rate: $\varepsilon = \frac{d \log wh}{d \log(1-\tau)}$. In practice, income taxes may distort choices beyond hours of work, such as training, effort, and fringe benefits. It is straightforward to incorporate such margins into the model by assuming that firms post job offers that specify H characteristics (or tasks), $\vec{h} = (h^1, \ldots, h^H)$, along with wage rates $\vec{w} = (w^1, \ldots, w^H)$ and workers have utility over-characteristics $\psi(h^1, \ldots, h^H)$. In such a model, the analysis that follows applies to the taxable income elasticity $\varepsilon = \frac{d \log \vec{w} \cdot \vec{h}}{d \log(1-\tau)}$ rather than the hours elasticity.

In the stylized models we consider here, the taxable income elasticity ε is the parameter relevant for analyzing tax policy (Feldstein 1999). In a more general union bargaining model with a finite labor demand elasticity, taxable income responses may be driven partly by wage and employment changes. For example, in Hansen's (1999) model of taxation with bargaining over wages and working hours, a higher marginal tax rate leads to lower wage rates, shorter working hours, and higher employment. Intuitively, when faced with an increase in tax rates, unions moderate their wage demands in exchange for a lower unemployment level. While the welfare implications of taxation would differ in such an environment, the three qualitative predictions derived below regarding the impact of frictions on observed responses to tax changes would still apply.

The elasticity ε is most commonly estimated using variation in tax rates from tax reforms (Blundell and MaCurdy 1999; Saez, Slemrod, and Giertz 2009). However, ε can also be identified from cross-sectional variation in tax rates using non-linear budget set methods (e.g. Hausman 1981). In particular, the amount of bunching observed at kinks identifies ε (Saez 2010). Let $B_{NL} = [F(\overline{\alpha}) - F(\underline{\alpha})]$ denote the fraction of type $s_i = NL$ individuals who choose $h_i = h_K$. Let $\tilde{g}_{NL}(h_K)$ denote the counterfactual density of hours in the absence of the tax change at the kink, which can be measured by the left limit of the density of the empirical hours distribution for type $s_i = NL$ individuals in this simple model. Under the approximation that the hours distribution g_{NL} is uniform around the kink, Saez (2010) shows that

$$\varepsilon \simeq \frac{B_{NL}(\tau_1, \tau_2) / \tilde{g}_{NL}(h_K)}{K \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)} = \frac{b_{NL}(\tau_1, \tau_2)}{K \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)}. \tag{6}$$

where $b_{NL} = B_{NL}/\tilde{g}_{NL}(h_K)$ denotes the fraction of type $s_i = NL$ individuals who bunch at the kink normalized by the counterfactual density. Intuitively, the fraction of individuals who stop working at $h_i = h_K$ hours because of the change in marginal tax rates is proportional to ε.

An important property of equations (5) and (6) is that the observed elasticity coincides with ε irrespective of the magnitude of the change in tax rates or the fraction of workers ζ affected by the tax change.[13] This result underlies microeconometric empirical studies of labor supply that use changes in taxes that affect subgroups of the population to identify ε. We now show that with search costs and hours constraints, observed elasticities vary with the size and scope of tax changes and no longer coincide with ε.

II.D. Special Case 2: Search Costs and Worker Responses

In this subsection, we analyze the impact of search costs on behavioral responses to taxation, abstracting from changes in the hours offered by firms. To isolate worker responses, we assume that the set of workers affected by the tax change has measure zero. When analyzing bunching at kinks, we assume that the fraction of agents who face the non-linear tax system is $\zeta = 0$; conversely, when analyzing tax reforms, we assume $\zeta = 1$. Under this assumption, the tax change has no impact on the equilibrium offer distribution $G(h)$ and only affects the treated workers' hours through changes in job search. To simplify notation, we assume that all workers face the same search costs $\phi_i = \phi$; the results below do not rely on this restriction.

Under these assumptions, a worker searches for a new job if his initial offer $h_i^0 \notin \left[\underline{h}_i, \overline{h}_i\right]$, where the thresholds are defined by the equations:

$$u\left(c_i(h_i^*), h_i^*\right) - u\left(c_i(\underline{h}_i), \underline{h}_i\right) = \phi \text{ with } \underline{h}_i < h_i^* \tag{7}$$

$$u\left(c_i(h_i^*), h_i^*\right) - u\left(c_i(\overline{h}_i), \overline{h}_i\right) = \phi \text{ with } \overline{h}_i > h_i^* \tag{8}$$

Workers who draw hours that fall within the region $\left[\underline{h}_i, \overline{h}_i\right]$ retain their initial offer because the utility gains from working h_i^* hours instead of h_i^0 hours are less than the cost of search ϕ. After the

13. We use the term "tax change" to refer both to changes in tax rates over time via reforms and changes in marginal tax rates at kinks within a given period.

search process is complete, there are two types of workers at each firm j: a point mass whose optimal labor supply $h_i^* = h_j$ is exactly that offered by the firm and a distribution of workers with optimal hours near but not equal to h_j.

Now consider how the mapping from the amount of bunching at kinks to ε in (6) is affected by search costs. Let $\hat{\varepsilon}(\tau_1, \tau_2) = \frac{B_{NL}(\tau_1,\tau_2)/\tilde{g}_{NL}(h_K)}{K \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)}$ denote the elasticity obtained by applying equation (6). We shall refer to $\hat{\varepsilon}$ as the "observed" elasticity from bunching at the kink. To understand the connection between $\hat{\varepsilon}$ and ε, first recall that in the frictionless model (where $\phi = 0$), workers locate at the kink if $\alpha_i \in [\underline{\alpha}, \overline{\alpha}]$. When $\phi > 0$, workers locate at the kink if $\alpha_i \in [\underline{\alpha}, \overline{\alpha}]$ *and* $h_i^0 \notin \left[\underline{h}_i, \overline{h}_i\right]$.[14] As a result, the observed elasticity $\hat{\varepsilon}$ is smaller than the structural elasticity ε. As the size of the tax change at the kink increases (τ_1 falls or τ_2 rises), the set of workers with $\alpha_i \in [\underline{\alpha}(\tau_1, \tau_2), \overline{\alpha}(\tau_1, \tau_2)]$ who pay the search cost to locate at the kink expands:

$$\frac{\partial[\overline{h}_i - \underline{h}_i]}{\partial \tau_2} < 0 \text{ and } \frac{\partial[\overline{h}_i - \underline{h}_i]}{\partial \tau_1} > 0. \tag{9}$$

Because the equilibrium hours distribution $G(h)$ is not affected by τ_1 and τ_2 when $\zeta = 0$, it follows immediately that $\hat{\varepsilon}$ rises with $\tau_2 - \tau_1$. As $\tau_1 \to -\infty$ and $\tau_2 \to \infty$, the inaction region $\left[\underline{h}_i, \overline{h}_i\right]$ collapses to h_K for agents with $\alpha_i \in [\underline{\alpha}, \overline{\alpha}]$ and $\hat{\varepsilon} \to \varepsilon$. Larger kinks generate larger observed elasticities because the utility costs of ignoring a kink increase with its size. Figure I illustrates this intuition using indifference curves in consumption-labor space for an agent who would optimally set hours at h_K. The thresholds $\left[\underline{h}_i, \overline{h}_i\right]$ are where the budget constraint crosses the indifference curve that yields utility ϕ units less than the maximal utility U^*. Now suppose τ_2 increases, moving the upper budget segment from the solid line to the dashed line. Then the upper bound $\overline{h}_i$ decreases, which in turn increases $\hat{\varepsilon}$. This is because the utility loss from supplying hours above the kink rises with τ_2, as one earns

14. Workers who draw $h_i^0 \in \left[\underline{h}_i, \overline{h}_i\right]$ do not contribute to the point mass at the kink because $G(h)$ is smooth when $\zeta = 0$. Therefore, among type $s_i = NL$ workers, the set who draw an initial hours offer $h_i^0 = K/w$ has measure zero. $G(h)$ is smooth in this case because the distribution of tastes $F(\alpha)$ is smooth and the set of agents who face a smooth (linear) tax schedule has measure 1.

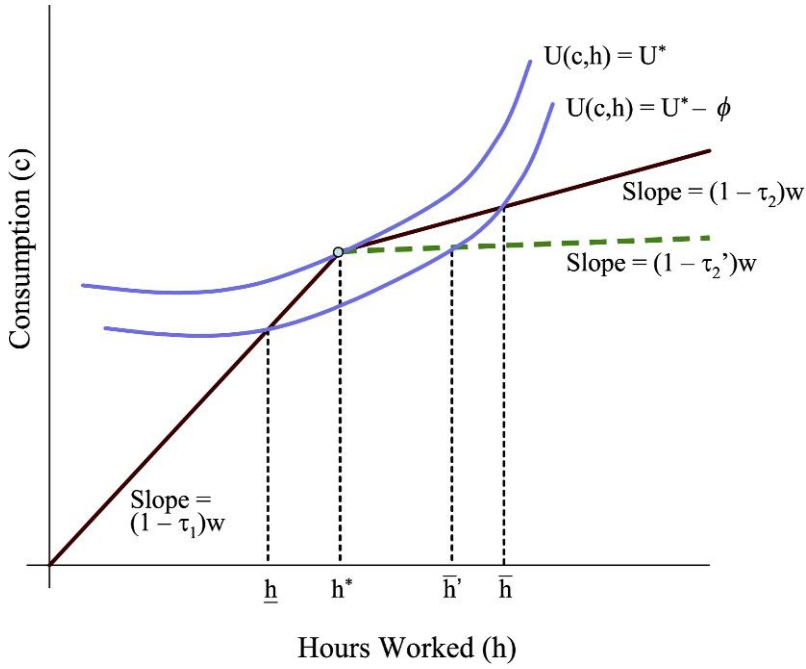

FIGURE I
Bunching at Kinks with Search Costs

This figure illustrates how search costs affect bunching at kinks. The two-bracket tax system creates the kinked budget set shown in dark gray. The worker's indifference curves are shown by the light gray isoquants. This worker's optimal labor supply is to set $h^*=h_K$, placing him at the kink. The lower indifference curve shows the optimal utility minus the search cost ϕ. If the workers draws an initial hours offer between $\overline{h}$ and $\underline{h}$, he will not pay ϕ to relocate to the kink. As the tax change at the bracket cutoff increases in magnitude (shown by the dashed line), the inaction region shrinks to $(\underline{h}, \overline{h}')$, leading to a larger observed elasticity from bunching.

less for this extra effort. These results lead to our first testable prediction:

PREDICTION 1: When workers face search costs, the observed elasticity from bunching rises with the size of the tax change and converges to ε as the size of the tax change grows:

$$\partial\hat{\varepsilon}/\partial\tau_2 > 0,\ \partial\hat{\varepsilon}/\partial\tau_1 < 0, \text{ and } \lim_{(\tau_2-\tau_1)\to\infty} \hat{\varepsilon} = \varepsilon \tag{10}$$

We derive an analogous prediction for observed elasticities from tax reforms in Online Appendix A. Tax reforms generate observed elasticities $\hat{\varepsilon} = \frac{d \log h}{d \log(1-\tau)}$ that differ from ε; as the size of the tax reform grows, $\hat{\varepsilon} \to \varepsilon$. The intuition for this result is very similar to that for bunching: many workers will not pay the search cost to find a job that requires fewer hours following a tax increase, attenuating $\hat{\varepsilon}$. However, unlike in the case of bunching, observed elasticities from tax reforms need not always be smaller than ε. For example, if workers are close to the edge of their inaction regions prior to the reform, a small tax change could lead to large adjustments, generating $\hat{\varepsilon} > \varepsilon$. Hence, observing that elasticities rise with the size of tax reforms is sufficient, but not necessary, to infer that search costs affect observed elasticities.

Non-Constant Elasticities. If the utility function is not isoelastic, one may observe an elasticity $\hat{\varepsilon}$ that increases with the size of the tax change even without search costs. We can distinguish search costs from variable elasticities by comparing the effects of several small tax changes with the effects of a larger change that spans the smaller changes. In Online Appendix A, we show that with an arbitrary utility $u(c,l)$ and tax rates $\tau_1 < \tau_2 < \tau_3$, the amount of bunching at two smaller kinks is equal to the bunching created at a single larger kink in the frictionless case ($\phi = 0$):

$$B_{NL}(\tau_1, \tau_3) = B_{NL}(\tau_1, \tau_2) + B_{NL}(\tau_2, \tau_3).$$

This is because the amount of bunching increases linearly with the size of the kink without search costs, as shown in (6). In contrast, when $\phi > 0$,

$$B_{NL}(\tau_1, \tau_3) > B_{NL}(\tau_1, \tau_2) + B_{NL}(\tau_2, \tau_3).$$

Intuitively, agents are more likely to pay the fixed search cost ϕ to relocate to the bigger kink, and thus it generates more bunching and a larger observed elasticity than the two smaller kinks together. A similar result applies to tax reforms: the observed effect of two small tax reforms, each starting from a steady state, differs from the effect of one large reform only when $\phi > 0$. We exploit these results to show that the differences in observed elasticities we document in our empirical analysis are driven by search costs rather than changes in the local elasticity.

Micro vs. Macro Elasticities. Search costs lead to a divergence between the elasticities observed from micro studies of tax reforms or bunching and the elasticities relevant for macroeconomic comparisons. In particular, the structural elasticity ε determines the steady-state effect of variation in tax policies across economies on aggregate labor supply even with search costs. To see this, consider two economies with different linear tax rates, τ and τ', for workers with $s_i = L$. To abstract from firm responses to this tax variation, assume that the set of individuals facing the linear tax has measure zero ($\zeta = 1$); we show that the same result holds with firm responses in the next subsection. We define the observed macro elasticity as the effect of this difference in tax rates on hours of work:

$$\widehat{\varepsilon}_{\text{MAC}} = \frac{\mathbb{E}\log h_i(\tau') - \mathbb{E}\log h_i(\tau)}{\log(1-\tau') - \log(1-\tau)}$$

For workers who pay the search cost to choose optimal hours, the difference in hours between the two economies is

$$\log h_i^*(\tau') - \log h_i^*(\tau) = \varepsilon \cdot (\log(1-\tau') - \log(1-\tau))$$

Workers who retain their original hours draw h_i^0 have average work hours of $\int_{\underline{h}_i}^{\overline{h}_i} h dG(h)$. Under a quadratic approximation to utility, the movement in the inaction region is also determined by ε:

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

Under the approximation that the offer distribution $G(h)$ is uniform between $\underline{h}_i$ and $\overline{h}_i$,

$$\mathbb{E}\log h_i(\tau') - \mathbb{E}\log h_i(\tau) \simeq \varepsilon \cdot (\log(1-\tau') - \log(1-\tau))$$

It follows that $\widehat{\varepsilon}_{\text{MAC}} \simeq \varepsilon$: the macro elasticity approximately equals the structural elasticity regardless of the search cost ϕ.

The critical difference between micro and macro elasticities is that the former are identified from a worker's decision to switch jobs *ex-post* because of tax incentives, whereas the latter are identified from differences in *ex-ante* job search behavior. Search costs reduce workers' propensity to fine tune their labor supply choices by bunching at kinks or responding to tax reforms because the costs of deviating from optima are second-order. But workers

search for jobs with fewer hours to begin with in an economy with higher tax rates. Consequently, a tax reform or a kink that changes the marginal rate from τ to τ' generates a smaller observed elasticity than the *same* "macro" variation in tax rates of τ vs. τ' across economies.

II.E. Special Case 3: Hours Constraints and Firm Responses

We now show how changes in hours constraints affect observed responses to tax changes. To highlight the importance of aggregate bunching and obtain analytical results, we consider a different special case of the model. First, we assume $\zeta \in (0, 1)$, so that there is a positive measure of workers affected by both tax systems. Second, we assume that at each level of α_i, a fraction δ of workers face no search costs ($\phi_i = 0$) and the remaining workers cannot search at all ($\phi_i = \infty$).

In this special case, workers' search decisions are simple: those with $\phi_i = 0$ choose $h_i = h_i^*$ and those with $\phi_i = \infty$ have $h_i = h_i^0$, their initial hours draw. As a result, the equilibrium distribution of job offers $G(h)$ coincides with the distribution of optimal hours choices, $G^*(h)$. The reason is that the search process $\mathcal{F}$ maps a distribution of offers to $\mathcal{F}(G) = \delta G^* + (1 - \delta) G$, and hence G^* is the only fixed point of $\mathcal{F}$. Intuitively, workers with $\phi_i = 0$ always choose their optimal hours, and so the only offer distribution that is a fixed point for them is G^*. As any offer distribution is a fixed point for the $\phi_i = \infty$ group, G^* must be the aggregate hours distribution in equilibrium. This result illustrates that hours constraints are determined by workers' aggregate tax preferences in equilibrium.

To see how the endogenous determination of hours constraints affects elasticity estimates, consider the observed elasticity from bunching for the workers who face the non-linear tax ($s_i = NL$). Let $B^*_{NL}(\tau_1, \tau_2)$ denote the total level of bunching that one would observe in the frictionless model ($\delta = 1$) for these workers. With search costs ($\delta < 1$), the observed amount of bunching for workers with $s_i = NL$ is:

$$B_{NL} = \delta B^*_{NL} + (1 - \delta)\zeta B^*_{NL}$$

The two terms in this expression represent two distinct sources of bunching. The first term arises from workers who choose $h_i = h_i^* = h_K$ because they face no search costs. The second term arises from the workers who set $h_i = h_i^0 = h_K$ because they face infinite search

costs. Because the aggregate distribution of hours coincides with the optimal aggregate distribution, a fraction ζB^*_{NL} of the equilibrium job offers have hours of h_K. We label the first component of bunching ($B^I_{NL} = \delta B^*_{NL}$) "individual bunching" because it arises from individuals' choices to locate at the kink via job search.[15] We label the second component ($B^A_{NL} = (1-\delta)\zeta B^*_{NL}$) "aggregate bunching" because it arises from the aggregation of workers' preferences by either unions or firms.

The signature of aggregate bunching is that it generates bunching even amongst workers who have no incentive to locate at the kink. Consider workers with $s_i = L$, who face a linear tax schedule and experience no change in marginal tax rates at h_K. Because of the interaction of hours constraints with search costs, these workers also bunch at the kink via the aggregate bunching channel. These workers draw $h^0_i = h_K$ with probability ζB^*_{NL} and are forced to retain that offer if $\phi_i = \infty$. The amount of bunching observed for workers with $s_i = L$ is therefore $B_L = (1-\delta)\zeta B^*_{NL} = B^A_{NL}$. This equivalence between B_L and B^A_{NL} is useful empirically because we cannot measure B^A_{NL} directly (as we do not observe search behavior), but we can measure B_L since we do observe workers' tax schedules. Intuitively, any bunching among those who do not face a kink must represent aggregate bunching.

The observed elasticity from bunching for workers with $s_i = NL$ is:

$$\hat{\varepsilon} = \frac{B_{NL}(\tau_1, \tau_2)/\tilde{g}^*_{NL}(h_K)}{K \ln\left(\frac{1-\tau_1}{1-\tau_2}\right)} = \delta\varepsilon + (1-\delta)\zeta\varepsilon < \varepsilon$$

The observed elasticity is smaller than the structural elasticity because search costs prevent some workers who would like to be at the kink from moving there.[16] The observed elasticity rises with

15. A fraction $(B^*_{NL})^2$ of workers with $\phi_i = 0$ and $h^*_i = h_K$ draw the $h^0_i = h_K$ to begin with and are therefore indifferent between retaining h^0_i and searching for their optimal job. To simplify notation, we classify these workers as "individual bunchers" by assuming that they choose to search for a new job.

16. In this special case, the *total* amount of bunching including all workers (both L and NL) equals the amount of bunching in the frictionless case ($\delta = 0$) because $G(h) = G^*(h)$. However, the composition of those at the kink differs when $\delta > 0$: some of those who bunch face the linear tax. This is why $\hat{\varepsilon} < \varepsilon$ for workers of type NL. In the general model where workers face finite adjustment costs, $G(h) \neq G^*(h)$ and total bunching no longer coincides with that in the frictionless case.

the scope of the kink ζ – the fraction of workers in the economy who face the non-linear tax schedule. When more workers face a change in tax incentives at an earnings level of K, firms are compelled to offer more jobs in equilibrium at h_K hours to cater to aggregate preferences. Thus a kink that affects more workers generates more aggregate bunching B^A_{NL} and thereby leads to more total bunching and a larger observed elasticity $\hat{\varepsilon}$.

As the scope of the kink approaches $\zeta = 1$, $B_{NL} \to B^*_{NL}$ and $\hat{\varepsilon} \to \varepsilon$ in this special case. Conversely, as ζ approaches 0, B^A_{NL} converges to 0 because firms only cater to aggregate preferences. It follows that the bunching observed at kinks that affect few workers in the economy constitutes a pure measure of individual bunching:

$$\lim_{\zeta \to 0} B_{NL} = B^I_{NL} \tag{11}$$

This equivalence between $\lim_{\zeta \to 0} B_{NL}$ and B^I_{NL} is also useful empirically because we cannot directly observe B^I_{NL}, but can observe $\lim_{\zeta \to 0} B_{NL}$ by studying bunching at kinks that apply to few workers.[17] These results lead to our second testable prediction.

PREDICTION 2: Search costs interact with hours constraints to generate aggregate bunching. Aggregate bunching and the observed elasticity rise with the fraction of workers who face the kink:

$$B^A_{NL} = B_L > 0 \text{ iff } \zeta > 0 \tag{12}$$

$$\frac{\partial B^A_{NL}}{\partial \zeta} > 0 \text{ and } \frac{\partial \hat{\varepsilon}}{\partial \zeta} > 0.$$

The source of aggregate bunching is that the distribution of jobs offered in equilibrium reflects the aggregation of workers' tax preferences. Therefore, in occupations where workers are more tax elastic, one should observe a higher level of both individual and aggregate bunching. To see this, consider the Q-sector extension of the model described above. The amount of individual bunching in occupation q is $B^I_{NL,q} = \delta \zeta B^*_{NL,q}$ and the amount of aggregate bunching is $B^A_{NL,q} = (1 - \delta)\zeta B^*_{NL,q}$. As the structural elasticity ε_q increases, the fraction of workers who would optimally locate at the kink ($B^*_{NL,q}$) increases, increasing both $B^I_{NL,q}$

17. This is why the bunching in special case 2 above (where $\zeta = 0$) is driven purely by individual search behavior rather than aggregate responses.

and $B^A_{NL,q}$ because δ and ζ are constant.[18] This leads to our third and final prediction.

PREDICTION 3: The amount of aggregate bunching and individual bunching are positively correlated across occupations:

$$cov\left(B^I_{NL,q}, B^A_{NL,q}\right) > 0 \tag{13}$$

Online Appendix A presents analogs of predictions 2 and 3 for observed elasticities from tax reforms.

Micro vs. Macro Elasticities. The structural elasticity ε continues to determine the macro elasticity with firm responses. Consider again the two economies with different linear tax rates, τ and τ', for workers of type $s_i = L$. But now assume that all workers face the linear tax ($\zeta = 0$), so that firms respond to this tax variation. The results above imply that the difference in equilibrium hours across the two economies coincides with the difference in optimal hours. It follows immediately that the difference in average hours of work between the two economies is

$$\mathbb{E}\log h_i(\tau') - \mathbb{E}\log h_i(\tau) = \mathbb{E}\log h_i^*(\tau') - \mathbb{E}\log h_i^*(\tau) = \varepsilon\cdot(\log\tau' - \log\tau)$$

Hence, the observed macro elasticity equals the structural elasticity ($\widehat{\varepsilon}_{\text{MAC}} = \varepsilon$) even in the presence of coordinate responses to taxes. This result highlights a second reason that the macroeconomic effects of taxes could be larger than microeconometric estimates. Variation in tax rates across economies shifts the aggregate distribution of workers' preferences and thereby induces changes in the hours constraints set by firms. In contrast, tax reforms or kinks that affect a small subgroup of workers do not generate substantial changes in hours constraints.

We derived the three predictions in special cases because the general model with finite search costs and endogenous hours constraints is analytically intractable. In Chetty et al. (2009) we use numerical simulations to verify that the three predictions hold in the general case. The simulations also show that the macro elasticity is typically close to ε in the general model. We therefore proceed to test the predictions empirically and determine the extent

18. If workers could switch between sectors, this correlation result would be reinforced because more elastic workers would sort toward sectors with more aggregate bunching.

to which adjustment costs and hours constraints attenuate micro elasticity estimates in practice.

III. Institutional Background and Data

The Danish labor market is characterized by a combination of institutional regulation and flexibility, commonly termed "flexicurity." The vast majority of private sector jobs are covered by collective bargaining agreements, negotiated by unions and employer associations. The collective bargains set wages at the occupation level as a function of seniority, qualifications, degree of responsibility, etc. The contracts are typically negotiated at intervals of 2–4 years. Despite this relatively rigid bargaining structure, rates of job turnover are relatively high and the unemployment rate is relatively low. For example, Andersen and Svarer (2007) report that rates of job creation and job destruction for most sectors and the overall economy in Denmark are comparable to those in the U.S. The unemployment rate in 2000 in Denmark was 5.4%, among the lowest in Europe.

During the period we study (1994–2001), income was taxed using a three-bracket system. Figure IIa shows the tax schedule in 2000 in terms of Danish Kroner (DKr). Note that \$1 $\approx$ DKr 6. The marginal tax rate begins at approximately 45%, referred to as the "bottom tax."[19] At an income of DKr 164,300, a "middle tax" is levied in addition to the bottom tax. The net-of-tax wage rate falls by 11% at the point where the middle bracket begins. Finally, at incomes above DKr 267,600, individuals pay the "top tax" on top of the other taxes, bringing the marginal tax rate to approximately 63%. The net-of-tax wage rate falls by 30% at the point where the top bracket begins. Approximately 25% of wage earners pay the top tax during the period we study. The large jump in marginal tax rates in a central part of the income distribution makes the Danish tax system particularly interesting for our purposes.[20]

Figure IIb plots the movement in the top bracket cutoff across years in real and nominal terms. Danish tax law stipulates that

19. Individuals with incomes below DKr 33,000 are exempt from this bottom tax; in practice, virtually all wage earners earn more than this threshold.

20. Denmark also has a complex transfer system that affects incentives for low incomes (Kleven and Kreiner 2006). We do not model the transfer system here because transfer programs affect very few individuals' marginal incentives around the middle and top tax cutoffs that are the focus of our empirical analysis.

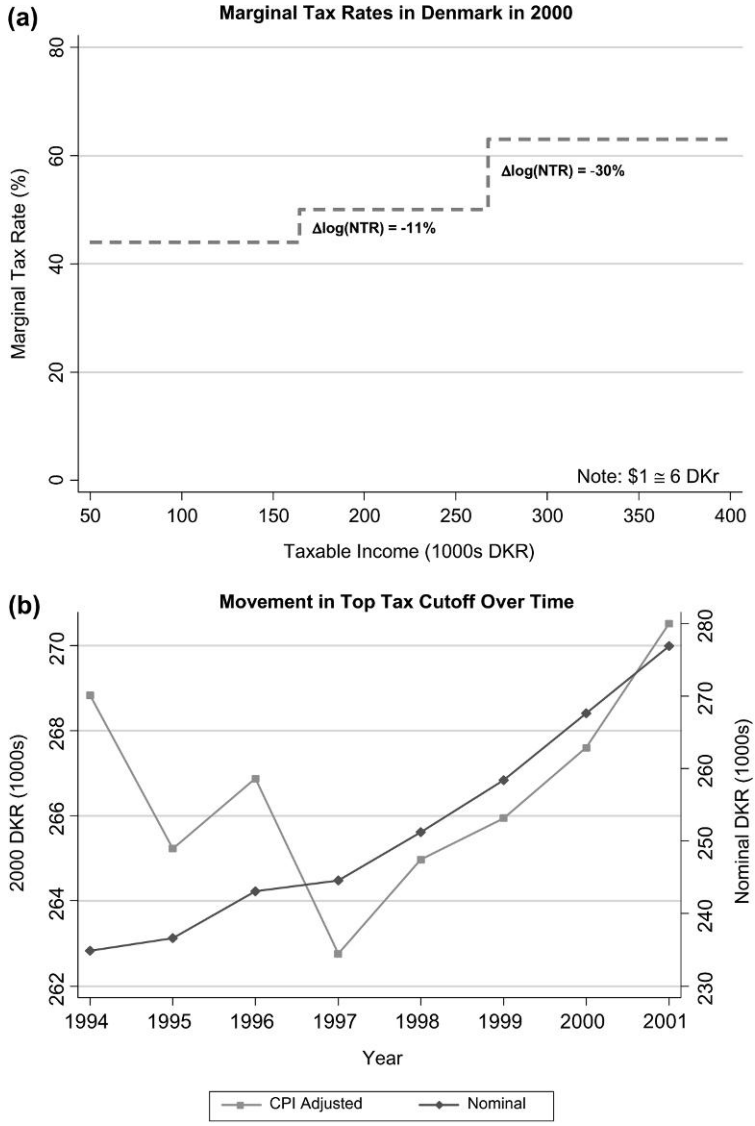

FIGURE II
The Danish Income Tax System

Panel (a) plots the marginal tax rate in 2000 vs. income for individuals living in Copenhagen, including the national tax, regional tax, and municipal tax. Panel (b) plots the level of taxable income above which earners must pay the top bracket national tax. The series in dark gray diamonds, plotted on the right *y*-axis, shows the nominal cutoff; the series in light gray squares, plotted on the left *y*-axis, shows the cutoff in real 2000 DKr.

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 mechanical, pre-determined movement of the cutoffs rules out potential concerns that the bracket cutoffs may be endogenously set as a function of labor market contracts. Over the period of study, inflation was between 1.8% and 2.9% per year. Because of the adjustment rule, the top bracket cutoff declines in real terms from 1994–1997 and then increases from 1998–2001.

In addition to the variation in tax rates across brackets, there were also some small tax reforms during the period we study. For example, in 1994 and 1995, there were two separate middle taxes that were consolidated into a single middle tax in subsequent years. Starting in 1999, net capital losses could not be deducted from the middle tax base and contributions to certain types of pensions could no longer be deducted from the top tax base. Finally, the middle and top tax bracket cutoffs change in real terms across years. These tax reforms generate changes in net-of-tax rates between −10% to +10% for certain subgroups, yielding several tax changes of small size and scope.

There are two tax bases relevant for our analysis: one for the top tax and one for the middle taxes. The top tax base depends almost entirely on individual income; the middle tax base is a function of household income. We study behavior at the individual level because our analysis focuses primarily on the top tax, but we account for joint aspects of the tax system when relevant (e.g. when studying the middle tax). 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.[21] Wage earnings, self-employment income, transfer payments, and gifts are all subject to both the middle and top income taxes. Most pension contributions are tax deductible and the marginal dollar of capital income is not subject to the top tax for most individuals. These features of the tax code create an incentive to shift earnings from labor income to capital income and pensions. See Ministry of Taxation (2002) for a more comprehensive description of the Danish tax system.

21. The Danish tax system includes a technical concept of "Taxable Income." Our use of the term "taxable income" does not refer to that technical concept.

Data. We merge several administrative registers provided by Statistics Denmark. The primary dataset is the tax register from 1994-2001, which contains panel data on wage earnings, self-employment income, pensions, capital income and deductions, spouse ID, and several other characteristics. The tax register contains records for more than 99.9% of individuals between the ages of 15–70 in the population. We merge the tax data with the Danish Integrated Database for Labor Market Research (IDA), which includes data on education, firm ID, occupation, labor market experience, and number of children for every person in Denmark. Additional details on the dataset and variable definitions are given in Online Appendix B.

Starting from the population dataset, we restrict attention to individuals who (1) are between the ages of 15 and 70 and (2) are wage earners, excluding the self-employed and pensioners.[22] These exclusions leave us with an analysis sample of 17.9 million observations of wage earners. Much of our analysis focuses on the subset of 6.8 million observations for wage earners that fall within 50,000 of the top tax cutoff. We also study the 1.8 million observations of self-employed individuals separately.

Table I presents summary statistics for the population of 15–70 year olds as a whole, all wage earners, the subset of wage earners within DKr 50,000 of the top tax cutoff, and self-employed individuals. The mean individual personal (non-capital) income in the population is DKr 180,213 ($30,000) for the population and DKr 227,359 ($38,000) for wage earners. Mean net capital income is negative because mortgage interest payments exceed capital income for most individuals. We define "net deductions" as deductions minus non-wage income (accounting for spousal deductions), or equivalently, wage earnings minus taxable income. Most wage earners have small net deductions (60% have deductions less than DKr 7,500 in magnitude), a fact that proves useful for our empirical analysis. The mean level of net deductions is negative because some individuals have substantial non-wage income.

We construct a tax simulator that calculates tax liabilities and marginal tax rates using these data. Given our focus on the top tax base, we compute marginal tax rates for individuals (i.e.,

22. The endogenous sample selection induced by dropping the self-employed does not spuriously generate bunching. There is significant bunching in the wage earnings distribution even in the full population: $b = 0.73$ in the full population vs. $b = 0.71$ for the subgroup of wage earners reported in Figure IIIa below.

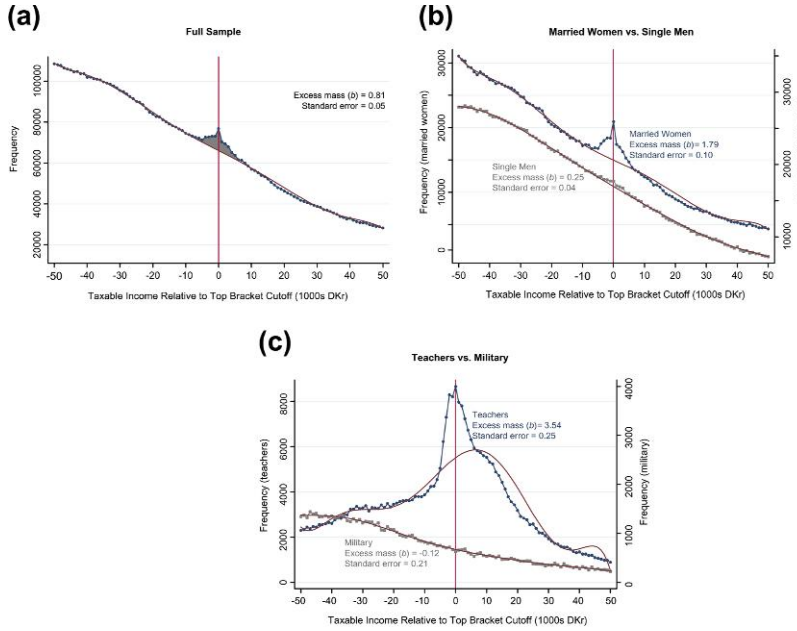

FIGURE III

Income Distributions around the Top Tax Cutoff for Wage Earners

These figures show the taxable income distribution around the top tax bracket cutoff (demarcated by the vertical line at 0) for wage earners between 1994–2001. The series shown in dots is a histogram of taxable income (as defined for the top tax base), relative to the top tax cutoff in the relevant year. Each point shows the number of observations 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 points DKr 7,500 or fewer from the cutoff, as in equation (15). In Panel (a) the full sample is considered. The shaded region is the estimated excess mass at the top bracket cutoff, which is 81% of the average height of the counterfactual distribution beneath. Panel (b) considers married women and single men. Panel (c) considers school teachers (ISCO 2331) and the military (ISCO 1013).

the change in tax liability for a given individual holding fixed spouse income) rather than households. We discuss below how this individual measure of marginal tax rates affects our analysis of bunching at the middle tax cutoff, which depends upon household income. Our tax simulator predicts actual tax liabilities within DKr 5 ($\simeq$ \$1) for 95% of the individuals in the population. Over the period we consider, top marginal tax rates were reduced slightly, and thus the simulated net-of-tax rate (holding

TABLE I
SUMMARY STATISTICS FOR THE DANISH POPULATION AND ESTIMATION SAMPLES, 1994–2001

	Population (1)	Wage Earners (2)	Wage Earners < DKr 50,000 from top tax cutoff (3)	Self employed (4)
Demographics:				
Age	40.91	39.17	41.43	46.02
Children	0.62	0.67	0.71	0.70
Labor market experience (yrs)	12.46	15.42	18.77	9.46
College education (%)	17.61	22.76	28.54	17.74
Female (%)	49.61	48.17	39.17	24.40
Married (%)	50.62	53.64	58.68	67.34
Income:				
Wage earnings	149,254	236,478	269,340	38,343
Other personal income	42,642	9,408	2,747	153,467
Total personal income	180,213	227,359	251,145	188,854
Net capital income	−10,672	−15,819	−19,570	−7,785
Deductions:				
Net deductions	−40,687	−13,151	−6,381	−31,996
|Net deductions| <7,500 (%)	43.25	59.36	69.11	23.84
|Net deductions-Pension kink| <7,500 (%)	2.03	2.72	2.96	5.07
Individual pension contributions	4,316	4,217	4,535	16,709
Employer pension contributions	7,584	12,065	13,131	2,123
Tax Payments:				
Pred. liability accurate within 5 DKr (%)	95.11	94.83	94.47	93.62
Pays the middle tax (%)	50.38	74.23	95.57	45.48
Pays the top tax (%)	18.06	25.87	33.53	23.61
2-year growth in NTR (%)	1.68	2.25	2.25	1.07
Std dev of 2-year growth in NTR (%)	4.50	4.95	4.95	6.80
Number of Obs.	30,492,819	17,866,090	6,788,235	1,846,064

Notes. Table entries are means unless otherwise noted. Column 1 is based on the full population of Denmark between ages 15-70 from 1994-2001. Column 2 includes all wage earners, the primary estimation sample. Column 3 includes only the subset of wage earners for whom |taxable income - top tax cutoff| <50,000, i.e. the individuals in Figure 3. Column 4 considers individuals who report positive self-employment income. All monetary values are in real 2000 Danish Kroner. Children are the number of children younger than 18 living with the individual. Personal income refers to all non-capital income. Net capital income refers to capital income minus payments such as mortgage interest. Net deductions refer to deductions from the top tax base such as individual pension contributions minus non-wage income such as taxable gifts. Net of tax rate is one minus the marginal tax rate predicted by our tax simulator.

fixed base-year characteristics) rises by 2.25% on average across two-year intervals.

IV. EMPIRICAL ANALYSIS

We begin by analyzing bunching at the top bracket cutoff, where net-of-tax wages fall by approximately 30%. In Figure IIIa

we plot the empirical distribution of taxable income for all wage earners in Denmark from 1994–2001. To construct this histogram, we first calculate the difference between the actual taxable income and the taxable income needed to reach the top tax bracket for each observation. We then group individuals into DKr 1,000 bins (−500 to 500, 500 to 1500, etc.) on this recentered taxable income variable. Finally, we plot the bin counts around the top bracket cutoff, demarcated by the gray vertical line at zero.

The figure shows that there is a spike around the top bracket cutoff in the otherwise smooth and monotonically declining income distribution. As shown in equation (6), the observed elasticity $\widehat{\varepsilon}$ implied by this bunching is proportional to $b(\tau_1, \tau_2)$, the excess mass relative to the density around the kink K. A complication in measuring b empirically is that the excess mass around K is diffuse rather than a point mass, presumably because it is difficult to control wage earnings perfectly. To measure b in the presence of such noise, we must estimate a counterfactual density – what the distribution would look like if there were no change in the tax rate at K. To do so, we first fit a polynomial to the counts plotted in the figure, *excluding* the data near the kink, by estimating a regression of the following form:

$$C_j = \sum_{i=0}^{q} \beta_i^0 \cdot (Z_j)^i + \sum_{i=-R}^{R} \gamma_i^0 \cdot \mathbf{1}[Z_j = i] + \varepsilon_j^0 \tag{14}$$

where C_j is the number of individuals in income bin j, Z_j is income relative to the kink in 1,000 Kroner intervals ($Z_j = \{-50, -49, .., 50\}$), q is the order of the polynomial, and R denotes the width of the excluded region around the kink (measured in DKr 1,000). Let B_N denote the excess number of individuals who locate at the kink. We define an initial estimate of the counterfactual distribution as the predicted values from (14) omitting the contribution of the dummies around the kink: $\widehat{C}_j^0 = \sum_{i=0}^{q} \widehat{\beta}_i^0 \cdot (Z_j)^i$. The excess number of individuals who locate near the kink relative to this counterfactual density is $\widehat{B}_N^0 = \sum_{j=-R}^{R} C_j - \widehat{C}_j^0 = \sum_{i=-R}^{R} \widehat{\gamma}_i^0$. This simple calculation overestimates B_N because it does not account for the fact that the additional individuals at the kink come from points to the right of the kink. That is, it does not satisfy the constraint that the area under the counterfactual must equal the area under the empirical distribution. To account for this problem, we shift the counterfactual distribution to the right of the kink upward until

it satisfies the integration constraint. In particular, we define the counterfactual distribution $\widehat{C}_j = \widehat{\beta}_i \cdot (Z_j)^i$ as the fitted values from the regression

$$(15)\quad C_j \cdot (1 + \mathbf{1}[j > R] \frac{\widehat{B}_N}{\sum_{j=R+1}^{\infty} C_j}) = \sum_{i=0}^{q} \beta_i \cdot (Z_j)^i + \sum_{i=-R}^{R} \gamma_i \cdot \mathbf{1}[Z_j = i] + \varepsilon_j$$

where $\widehat{B}_N = \sum_{j=-R}^{R} C_j - \widehat{C}_j = \sum_{i=-R}^{R} \widehat{\gamma}_i$ is the excess number of individuals at the kink implied by this counterfactual.[23] Finally, we define our empirical estimate of b as the excess mass around the kink relative to the average density of the counterfactual earnings distribution between $-R$ and R:

$$(16)\qquad \widehat{b} = \frac{\widehat{B}_N}{\sum_{j=-R}^{R} \widehat{C}_j / (2R + 1)}$$

The solid curve in the figure shows the counterfactual density $\{\widehat{C}_j\}$ predicted using this procedure with a seventh-degree polynomial ($q = 7$) and a window of DKr 15,000 centered around the kink ($R = 7$). The shaded region shows the estimated excess mass around the kink. With these parameters, we estimate $b = 0.81$ – the excess mass around the kink is 81% of the average height of the counterfactual distribution within DKr 7,500 of the kink. The qualitative results we report below are not sensitive to changes in q and R or the way in which we correct the counterfactual to satisfy the integration constraint. The reason is that the differences we document in observed elasticities are much larger than the changes induced by varying the specification of the counterfactual.

We calculate a standard error for $\widehat{b}$ using a parametric bootstrap procedure. We draw from the estimated vector of errors ξ_j in (15) with replacement to generate a new set of counts and apply the technique above to calculate a new estimate $\hat{b}^k$. We define the standard error of $\hat{b}$ as the standard deviation of the distribution of $\hat{b}^k$s. Since we observe the exact population distribution of taxable income, this standard error reflects error due to misspecification of the polynomial for the counterfactual income distribution rather

23. Because $\widehat{B}_N$ is a function of $\widetilde{\beta}_i$, the dependent variable in this regression depends upon the estimates of $\widetilde{\beta}_i$. We therefore estimate (15) by iteration, recomputing $\widehat{B}_N$ using the estimated $\widetilde{\beta}_i$ until we reach a fixed point. The bootstrapped standard errors that we report below adjust for this iterative estimation procedure.

than sampling error. The standard error associated with our estimate of b is 0.05. The null hypothesis that there is no excess mass at the kink relative to the counterfactual distribution is rejected with a t-statistic of 17.6, implying $p < 1 \times 10^{-9}$.

There is substantial heterogeneity across groups in the amount of bunching. Figure IIIb shows that excess mass at the kink is much larger for married women ($b = 1.79$) than for single men ($b = 0.25$), consistent with existing evidence that married women exhibit the highest labor supply elasticities.[24] Figure IIIc shows that there is also substantial heterogeneity across occupations: teachers exhibit substantial bunching around the kink ($b = 3.54$), whereas the military does not ($b = -0.12$, statistically insignificant).[25] We return to explore the sources of this heterogeneity in Section IV.B below.

The identification assumption underlying causal inference about the effect of taxes on earnings in the preceding analysis is that the income distribution would be smooth 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. Figure IV displays the distribution of taxable income in each year from 1994–2001 for all wage earners and for married women. The excess mass for both groups follows the movement in the top bracket cutoff very closely. In Figure V, we investigate whether the excess mass tracks tax changes, inflation, or average wage growth over time. We consider the period from 1997 to 2001, during which the top tax cutoff rises in real terms. Noting that the excess mass is located at the top tax cutoff in 1997, the figure shows three possibilities for its location in 2001: the 2001 top tax cutoff, the 1997 cutoff adjusted for inflation, and the 1997 cutoff adjusted for average wage growth in the economy. In both the full population of wage earners and the subgroup of married women, the excess mass that was at the 1997 kink clearly moves to the 2001 kink rather than following inflation or average wage growth. The same pattern is observed during other periods when the top tax cutoff is declining in real terms (see

24. In principle, the bunching for married women could be exaggerated by wage payments from self-employed husbands seeking to reduce their tax liabilities. In practice, we find that the amount of bunching is virtually unchanged when we exclude households with at least one self-employed person from the sample.

25. Approximately 50% of wage earners in Denmark work in the public sector. We find slightly more bunching for those employed in the private sector ($b = 0.67$) than those in the public sector ($b = 0.5$).

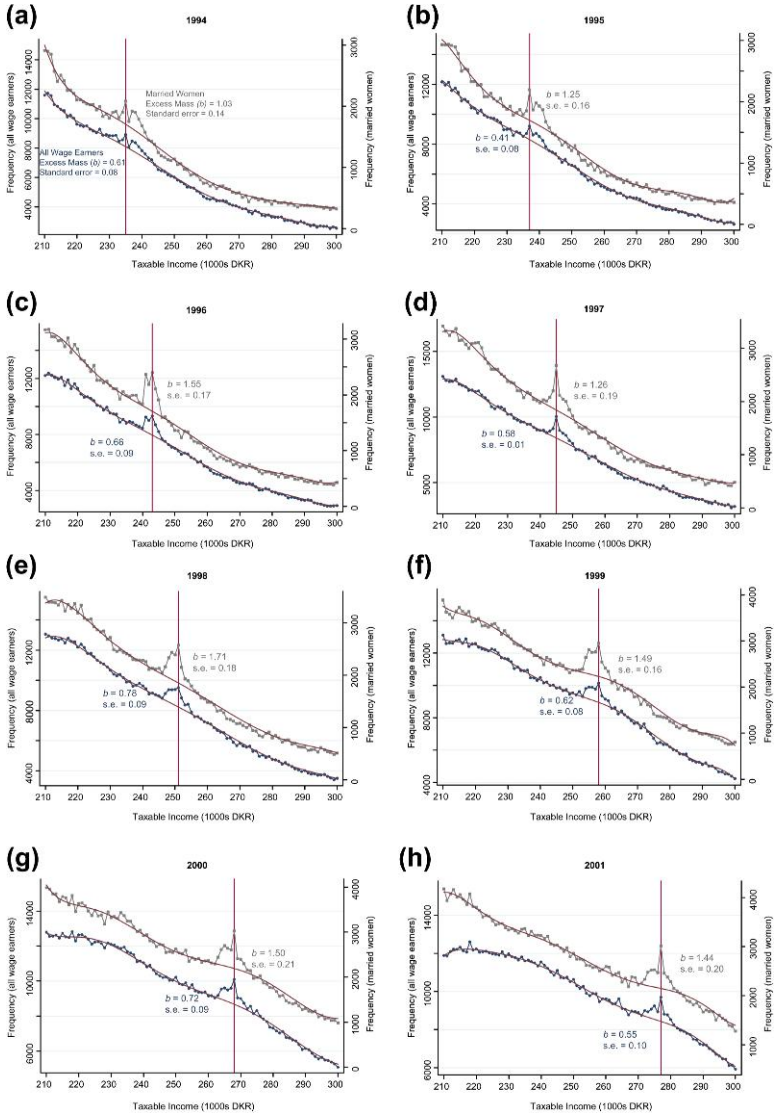

FIGURE IV
Income Distributions Around the Top Tax Cutoff, 1994–2001

These figures plot the empirical distribution of taxable income for wage earners and married female wage earners in each year from 1994–2001. In all panels, the upper distribution is for married women and the lower distribution is for all wage earners. The solid vertical lines mark the top tax bracket cutoff (in nominal DKr) in each year. The figure also shows the counterfactual distributions and excess masses, computed as in Figure IIIa.

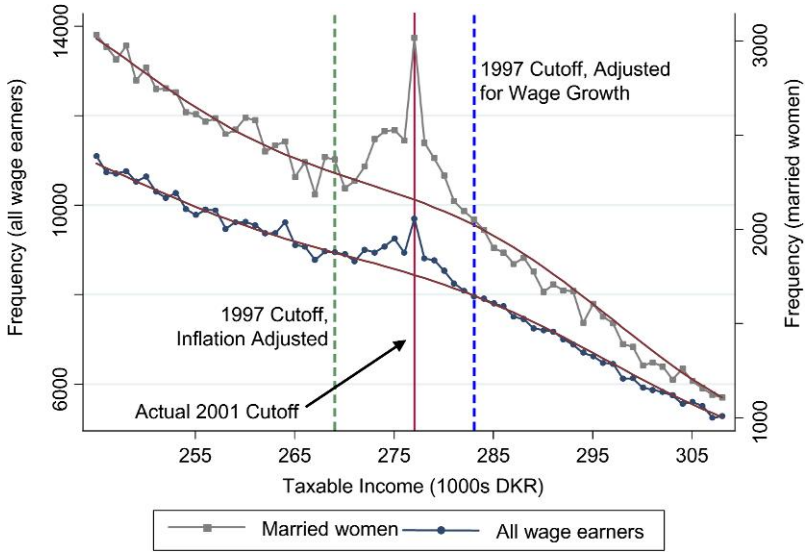

FIGURE V

Distinguishing Changes in Tax Incentives from Inflation and Wage Growth

This figure replicates the income distribution in Figure IV(h) for the year 2001, zooming in around the top tax bracket cutoff. The location of the bracket cutoff in 2001 is marked with the solid line. The dashed light gray line shows the level of the 1997 top bracket cutoff adjusted for inflation. The dashed dark gray line shows the 1997 bracket adjusted for average wage growth.

Figure A.1 in the Online Appendix). Earnings dynamics around the top tax bracket depart from prevailing inflation patterns and instead are aligned with changes in the tax system. We show that firm responses explain why the excess mass tracks the movement of the kink so closely despite frictions in Figure XI below.

Shifting vs. Real Responses. Individuals can obtain 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 forms of compensation. Our three theoretical predictions about how frictions affect observed taxable income elasticities hold regardless of what margins underlie changes in taxable income. Intuitively, if firms face technological constraints that limit the benefit packages workers can choose from, tax changes of larger size and scope will continue to produce larger taxable income elasticities. Nevertheless, it is useful to distinguish between these two behavioral responses because income shifting and "real" changes in labor supply have

different normative implications (Slemrod and Yitzhaki 2002; Chetty 2009).

There are two channels through which individuals can change their reported taxable income without changing labor supply: evasion and avoidance. Kleven et al. (2010) study audited Danish tax records and find that there is virtually no tax evasion in wage earnings because of third-party reporting by firms. We find that there is substantial bunching ($b = 0.68$) even in wage earnings (see Figure A.2). We therefore conclude that the bunching we observe is not driven by evasion.

The second and more important income shifting channel is legal tax avoidance. The simplest method of reducing current tax liabilities is to contribute to tax-deductible pension accounts. We investigate the extent of such shifting by adding employer and employee pension contributions back to taxable income. We find that the distribution of this broader measure of compensation still exhibits substantial bunching relative to the statutory top tax bracket cutoff that would apply to individuals with *zero* pension contributions, rejecting the hypothesis that all of the bunching observed in taxable income is driven by shifts to pensions (see Figure A.2).We conclude that pension shifting is responsible for only a small amount of the bunching in taxable income we observe at the top tax cutoff. The relatively small amount of pension shifting is likely driven by the generosity of Denmark's social security programs. An analogous exercise shows that shifting into capital income, which is untaxed in the top tax base, is responsible for virtually none of the bunching at the top kink.

Although the behavioral responses at the top tax cutoff do not appear to be driven by any observable method of income shifting, 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 office amenities when individuals cross into the top tax bracket. We also cannot rule out intertemporal shifting of wage earnings to avoid paying the top tax. The only way to definitively rule out such responses is to examine changes in hours worked directly. Unfortunately, our dataset does not contain information on hours of work. Nevertheless, we believe that most of the observed bunching in taxable income reflects "real" distortions in behavior that have efficiency costs. Few salaried workers at the 75th percentile of the income distribution have the ability to shift income into other forms of

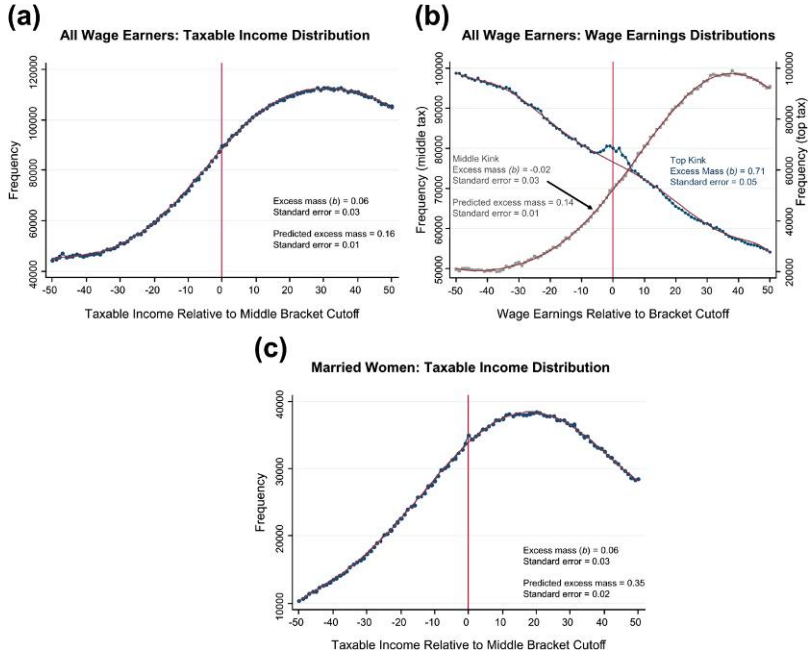

FIGURE VI
Income Distributions around the Middle Tax Cutoff

Panels (a) and (c) plot the empirical distributions of taxable income (as defined for the middle tax base) around the middle tax cutoff, where net-of-tax wage rates fall by approximately 10%. Panel (b) plots the distribution of wage earnings around the middle and top tax cutoffs. Panels (a) and (b) include all wage earners, while panel (c) includes only married female wage earners. All panels show counterfactual distributions and excess masses, computed as in Figure IIIa. Each panel also reports the amount of bunching predicted if the elasticity were the same as that estimated from the amount of bunching at the top bracket cutoff for the corresponding income measure and subgroup.

compensation or across time (Slemrod 1995; Goolsbee 2000). Moreover, even if compensation is distorted toward office amenities instead of wages, the marginal efficiency cost of such distortions equals the marginal efficiency cost of changes in hours of work (Feldstein 1999).

IV.A. Prediction 1: Size of Tax Changes

We now test the first prediction by comparing the amount of bunching at the top tax kink with bunching at smaller kinks and observed elasticities from small tax reforms. Figure VI

shows the distributions of taxable income around the middle tax cutoff, where the net-of-tax rate falls by approximately 10%.[26] Figure VIa shows that there is virtually no bunching at the middle tax cutoff ($b = 0.06$) in taxable income for the full population of wage earners. Moreover, the estimated excess mass at the middle tax converges to zero as the degree of the polynomial is increased, whereas the estimated excess mass at the top kink is not sensitive to the degree of the polynomial. Because the definitions of "taxable income" differ for the top and middle tax bases, Figure VIb plots the distribution of wage earnings around both kinks. Consistent with Figure VIa, there is significantly more bunching at the top kink than the middle kink in wage earnings. Figure VIc shows that the amount of bunching remains small and statistically insignificant even for the subsample of married women, who exhibit substantial bunching at the top kink as shown in Figure IIIb.

Note that smaller kinks should generate less bunching even in the frictionless model, simply because the change in incentives is smaller. We therefore compare the 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. In all cases, the amount of bunching observed in the empirical distribution at the middle kink is significantly less than what would be predicted by the frictionless model. For example, the frictionless model predicts $b = 0.16$ at the middle kink for all wage earners (Figure VIa). The null hypothesis that the predicted excess mass equals the actual excess mass at the middle kink can be rejected with $p < 0.01$.

Next, we estimate observed elasticities using changes in marginal rates by legislated reforms. As described in Section III, several small tax reforms in Denmark between 1994 and 2001 created changes in net-of-tax rates of between -10% and $+10\%$. These reforms generate differential changes in net-of-tax rates across income groups, motivating a difference-in-difference research design. Let $\Delta \log y_{i,t} = \log y_{i,t} - \log y_{i,t-2}$ denote the log change in wage earnings from period $t-2$ to t and $\Delta \log(1-MTR_{i,t})$ the log change in net-of-tax rates over the same period. Following Gruber and Saez (2002), we estimate the following regression

26. In 1994 and 1995, the tax system includes an additional "upper middle tax." Figure VI only considers the lower middle tax in these years, but there is no bunching at the upper middle tax cutoff either.

specification using two-stage-least-squares:

$$(17) \quad \Delta \log y_{i,t} = \alpha + \beta \Delta \log(1 - MTR_{i,t}) + f(y_{i,t-2}) + \gamma X_{i,t-2} + \varepsilon_{i,t},$$

instrumenting for $\Delta \log(1 - MTR_{i,t})$ with $\Delta \log(1 - MTR_{i,t}^{\text{sim}})$, the simulated change in net-of-tax rates holding the individual's income and other characteristics fixed at their year $t-2$ levels. The function $f(y_{i,t-2})$ is a 10-piece linear spline in base year wage earnings and the vector $X_{i,t-2}$ is a set of base year controls that we vary across specifications. First-stage regressions of $\Delta \log(1 - MTR_{i,t})$ on $\Delta \log(1 - MTR_{i,t}^{\text{sim}})$ have coefficients of approximately 0.6 with t-statistics exceeding 600.

Table II reports TSLS estimates from several variants of (17). In column (1), we estimate (17) on the full population of wage earners with the following controls: the 10-piece wage earnings spline, a 10-piece spline in total personal income and age and year fixed effects. The estimated elasticity $\widehat{\varepsilon}$ is very close to 0, and the upper bound of the 95% CI is $\widehat{\varepsilon} = 0.004$. Column (2) adds a 10-piece capital income spline, gender and marital status dummies, and occupation and region 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 income growth. Column (3) considers the subgroup of married women using the baseline specification in column (1). The observed elasticity in response to small tax changes remains near 0 for married women despite the fact that they exhibit substantial bunching at the large top tax kink, as shown in Figure IIIb. In column (4), we further restrict the sample to married women who are professionals and have above-median (more than 19 years) labor market experience. This subgroup also does not react significantly to small tax reforms, yet it exhibits substantial bunching at the top kink (b=4.50, implying $\widehat{\varepsilon} = 0.06$).

In sum, our analysis confirms that larger tax changes produce larger observed elasticities. However, the elasticity implied by the frictionless model remains very small even at the largest kink. The observed elasticity from bunching at the 30% kink is $\widehat{\varepsilon} \simeq 0.01$ for all wage earners and $\widehat{\varepsilon} \simeq 0.02$ for married women. We believe that these elasticity estimates remain substantially attenuated relative to ε because the utility loss from ignoring the 30% change in tax rates at the top kink is only around 2% of consumption given a structural elasticity of $\varepsilon = 0.5$ (Chetty 2011).

TABLE II
OBSERVED ELASTICITY ESTIMATES USING SMALL TAX REFORMS

	Dependent Variable: Log Change in Wage Earnings					
Subgroup:	All Wage Earners		Married Females	High-Experience Married Female Professionals	Wage Earners 100-300K	Wage Earners > 200K
	(1)	(2)	(3)	(4)	(5)	(6)
Log change in net-of-tax rate (Δ log (1–t))	−0.001 (0.003)	−0.004 (0.003)	0.006 (0.005)	0.000 (0.011)	−0.006 (0.003)	−0.001 (0.003)
Earnings $\times \Delta$ log (1–t)					0.005 (0.008)	
Labor income spline	x	x	x	x	x	x
Personal income spline	x	x	x	x	x	x
Year fixed effects	x	x	x	x	x	x
Age fixed effects	x	x	x	x	x	x
Region fixed effects		x			x	
Occupation fixed effects		x			x	
Capital income spline		x			x	
Gender, Marital status		x			x	
Observations	11,682,651	8,302,905	3,203,515	159,661	7,967,169	7,585,011

Notes. Standard errors clustered by individual reported in parentheses. Dependent variable in all specifications is two-year growth rate in real wage earnings. Independent variable of interest is two-year growth rate in net-of-tax rate, instrumented using two-year growth rate in simulated net-of-tax rate using base-year variables. Coefficients reported can be interpreted as observed wage earnings elasticities from tax reforms. All specifications include 10-piece wage earnings and total personal income splines as well as age and year fixed effects. Column 2 also includes a 10-piece capital income spline, gender and marital status indicators, and region and occupation fixed effects. Occupation fixed effects are available only for a subset of years and observations. Column 4 restricts attention to married female professionals with more than 19 years of labor market experience. Column 5 interacts the log change in net of tax rate with the difference between wage earnings and the top tax cutoff (measured in DKr 100,000) to test whether the taxable income elasticity varies by income level. This specification restricts the sample to individuals with wage earnings in the base year between DKr 100,000 and 300,000. Column 6 considers individuals with more than DKr 200,000 of wage earnings in the base year.

Search Costs vs. Non-Constant Elasticities. If $\varepsilon(\tau, z)$ varies with τ or z, the evidence that larger tax changes generate larger observed elasticities could potentially be explained by variation in ε rather than adjustment costs. In our application, the middle kinks are at incomes of DKr 130,000–177,900, while the top kinks are at incomes of DKr 234,900–276,900. If higher income individuals are more elastic, one would observe more bunching at the top kink even without frictions. We distinguish this explanation of our findings from frictions using three approaches.

First, we test whether taxable income elasticities differ by income by interacting $\Delta \log(1 - MTR_{i,t})$ with $y_{i,t-2}$ (re-centered around the top tax cutoff). Column (5) of Table II shows that this interaction effect is small and insignificant ($p = 0.52$), indicating that there is no significant heterogeneity in observed elasticities by income. As an alternative approach to assessing heterogeneity, we replicate the baseline specification in column (1) restricting the sample to individuals with wage earnings exceeding DKr 200,000. Column (6) shows that the estimated elasticity remains very close to zero, confirming that small tax changes do not generate significant behavioral responses even for individuals facing the top tax.

Second, we examine how the degree of bunching changes as the middle and top tax cutoffs move across years. In the latter years of our sample, the middle tax cutoff is higher in the income distribution, but the amount of bunching remains near zero (not shown). In contrast, bunching at the top kink remains substantial in all years (Figure IV).

As a third 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 50, 000 of the top kink in year t and within DKr 50, 000 of the middle kink in year $t + 2$. By studying these "switchers," we can effectively remove individual fixed effects when comparing responses to the middle and top kinks. We find that when near the top kink, these switchers exhibit substantial bunching ($b = 0.54$). However, just two years later, the same individuals show no excess propensity to bunch at the middle kink ($b = 0.06$) despite having earnings near that kink (see Figure A.3). The opposite pattern is observed for those moving from the middle to the top kink. We conclude that variation in observed elasticities is unlikely to explain the positive relationship between larger tax changes and larger observed elasticities.

Jointness of the Middle Tax Cutoff. As noted above, the Danish tax system has more elements of jointness at the middle kink than the top kink. In particular, spouses can transfer deductions between each other to minimize their middle tax liabilities, effectively making the middle tax a function of household income. Our individual-based measure of bunching at the middle tax is accurate if individuals make wage earnings decisions based on their own tax liabilities. However, our method could in principle understate the amount of bunching at the middle tax cutoff if spouses choose their earnings levels to minimize the tax burdens of the household as a whole rather than their own liability. As we explain in Online Appendix B, our method of computing bunching effectively computes the higher earner's distance to the kink based on the joint tax liability of the household rather than the individual. We find that bunching at the top tax cutoff remains significantly larger than at the middle tax cutoff for the subsample of individuals who are either the higher earner in a couple or are single (see Figure A.4). This result confirms that the differences in observed elasticities at the top and middle kinks shown in Figures III and IV are robust to the way in which we account for the jointness of the middle kink.[27]

Perceptions of the Middle vs. Top Cutoffs. What are the costs that workers face in responding to tax incentives? One possibility is the cost of paying attention to taxes (e.g. Chetty and Saez 2009). Figure A.5 reports the distribution of perceived middle and top tax cutoff obtained from an internet survey of 3,299 individuals who were members of a union representing public and financial sector employees (FTF-A).[28] The figure shows that knowledge of the top tax cutoff is better than the middle tax cutoff. The same qualitative pattern is exhibited across all education levels and occupations in the sample. These survey responses must be viewed as anecdotal evidence because the survey was administered only to members of FTF-A and because the response rate is low (11%). Nevertheless, this evidence is consistent with our finding that

27. A further concern is that it there may be differences in the costs of bunching at joint vs. individual kinks. For instance, jointness may allow the spouse with lower adjustment costs (e.g. the secondary earner) to choose a job that places the household at the kink. Such effects would work against finding more bunching at the top kink than the middle kink.

28. We thank Anders Frederikssen for making these data available to us.

observed elasticities are larger at the top kink than the middle kink.

IV.B. Prediction 2: Aggregate Bunching and Scope of Tax Changes

To test the second prediction, we begin by identifying a source of variation in the scope of kinks – the fraction of workers in the economy who face a given kink in the tax system. Recall that taxable income is the sum of wage earnings and non-wage income minus deductions. Deductions consist primarily of pension contributions. Non-wage income includes items such as alimony receipts, stipends, and unemployment benefits. Because of heterogeneity in non-wage income and deductions, the wage earnings required to reach the middle and top brackets vary across individuals.

Approximately 60% of wage earners have net deductions (deductions minus non-wage income) less than DKr 7,500 in magnitude (see Figure A.6). This is because most individuals in Denmark make no tax deductible pension contributions and earn only wage income. Thus, most individuals cross into the top tax bracket when their wage earnings exceed the top tax cutoff that applies to taxable income, which we term the "statutory" top tax cutoff. The distribution of deductions for the remaining 40% of individuals is diffuse, with one important exception. There is a mass point in the distribution of deductions at approximately DKr 33,000, which is driven by a cap on tax-deductible pension contributions. Individuals who make pensions contributions up the cap (approximately 2.7% of wage earners) reach the top tax bracket only when their wage earnings exceed the statutory top tax cutoff by DKr 33,000.

In this setting, the second prediction of our model consists of three parts: we should observe (1) significant aggregate bunching at the statutory top tax kink that applies to 60% of workers, (2) little aggregate bunching at the "pension kink" that applies to 2.7% of workers, and (3) more bunching for individuals with small deductions, as they have more common tax preferences. To test these hypotheses, we study wage earnings distributions at the occupation level because most wages are set through collective bargains at the occupation level in Denmark.

Aggregate bunching is easiest to see through case studies of occupations. Consider school teachers, who constitute approximately 3% of wage earners in Denmark and form one of

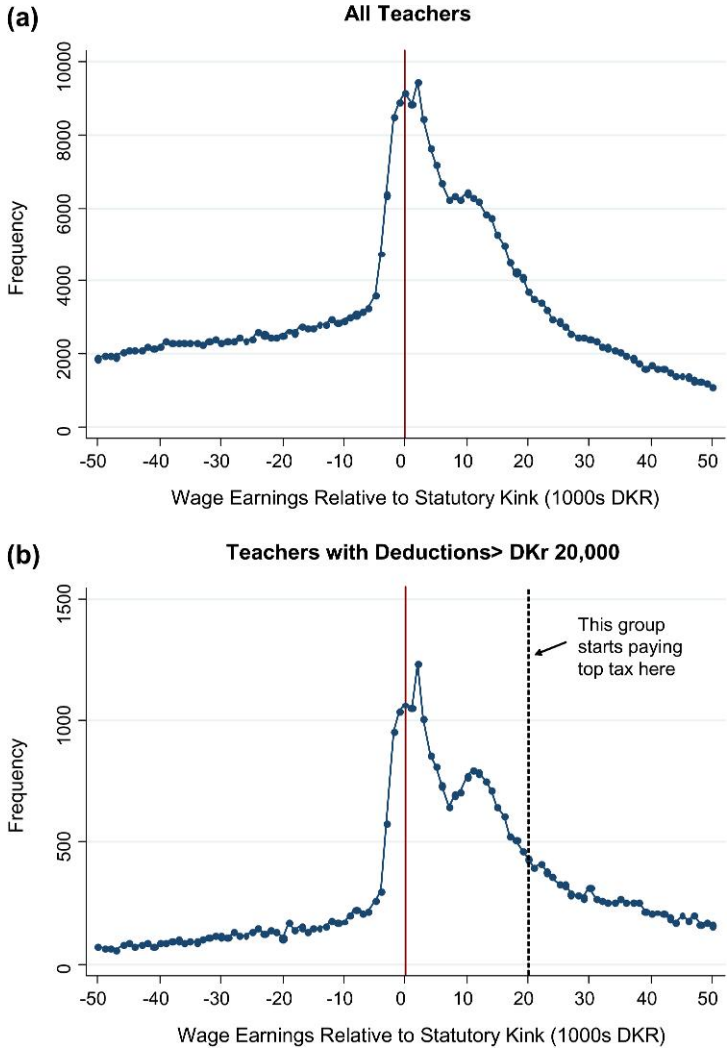

FIGURE VII
Teachers' Wage Earnings Distributions

These two figures plot the empirical distribution of wage earnings around the statutory top tax cutoff in 1994–2001 for (a) all teachers (ISCO 2331) and (b) teachers with net deductions greater than DKr 20,000.

the largest unions. Figure VIIa plots the distribution of wage earnings around the top tax bracket for teachers. There is very sharp bunching around the statutory top tax cutoff, consistent with the

sharp bunching in taxable income shown in Figure IIIc.[29] Intuitively, 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 the teachers union starts bargaining on other dimensions, such as lighter teaching loads or more vacations, rather than continue to push for wage increases beyond this point.

Figure VIIb plots the distribution of wage earnings (salaries) around the statutory top tax cutoff for teachers with net deductions *greater* than DKr 20,000. The individuals in this figure do not begin to pay the top tax on wage earnings until at least DKr 20,000 beyond the statutory top tax cutoff, and therefore experience no change in net-of-tax wages at the vertical line at zero. Yet the wage earnings distribution for these workers is extremely similar to the distribution for teachers as a whole, and exhibits sharp bunching at the *statutory* top tax cutoff. This is the signature of aggregate bunching: even individuals who are unaffected by a kink bunch there. In our model, those with deductions greater than DKr 20,000 effectively have type $s_i = L$ around the statutory kink; Figure VIIb shows that $b_L = b^A_{NL} > 0$. Intuitively, school districts offer a limited number of wage-hours packages in order to coordinate class schedules. Because of such technological constraints, teachers' contracts cater to the most common tax incentives in the population (i.e., those with small deductions).

There are similar patterns of aggregate bunching in many other occupations. We generalize from such case studies by analyzing the modes of the earnings distribution in each occupation, defined using four digit International Standard Classification of Occupations (ISCO) codes. We define the mode in each occupation-year cell as the DKr 5,000 wage earnings bin that has the largest number of workers. Figure VIII shows a histogram of these modes relative to the top tax bracket cutoff, excluding small occupation-years that have less than 7,000 workers (25% of the sample). The density of modes drops sharply at the top tax threshold. There are 20 modes within DKr 2000 of the top tax cutoff, but only 6 in

29. The smaller peak above the kink is driven by teachers in Copenhagen, who receive a cost-of-living adjustment of approximately DKr 15,000 over the base teacher's salary. The setting of salaries to place teachers outside Copenhagen – who account for 75% 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.

the adjacent bin from DKr 2,000 to DKr 6,000 above the kink. This drop in the frequency of modes across these two bins is larger than any other drop across two contiguous bins in the figure. Moreover, as the top tax cutoff rises over years, the distribution of modes shifts along with the cutoff (not shown). Hence, aggregate tax incentives – which are determined largely by the preferences of workers who face the statutory cutoff – shape the distribution of jobs offers.

Having established the prevalence of aggregate bunching at the most common kink, we test whether kinks that affect fewer workers generate less aggregate bunching. To do so, we exploit the "pension kink" described above. Figure IXa plots the distribution of wage earnings relative to the pension kink (shown by the vertical line at 0) for individuals who have deductions greater

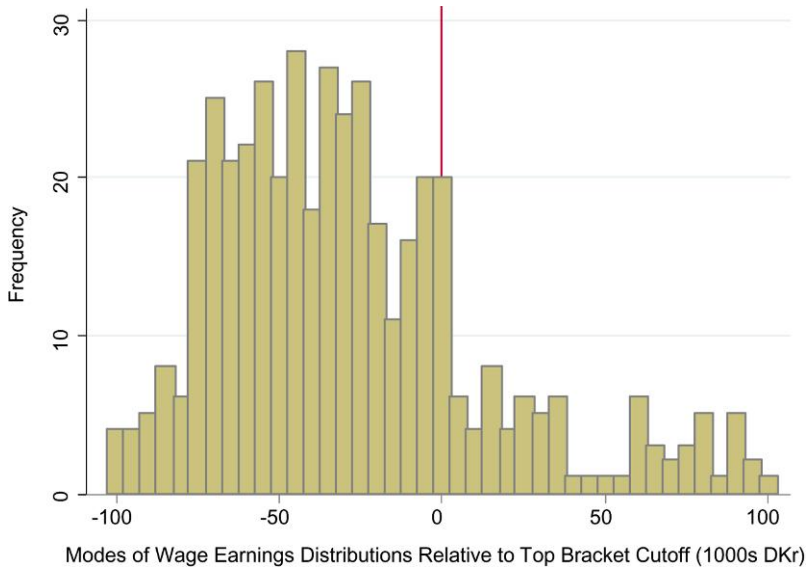

FIGURE VIII
Modes of Occupation-Level Wage Earnings Distributions

To construct this figure, we calculate the mode of the wage earnings distribution in each occupation-year cell, defined as the DKr 5,000 bin with the most individuals in that occupation-year. Occupations are defined by 4 digit ISCO codes. The figure shows a histogram of these modes, excluding occupations with fewer than 7000 workers.

than DKr 20,000. There is significant bunching in wage earnings at the top tax pension kink ($b = 0.70$).[30] To investigate whether this bunching is driven by aggregation of workers' tax preferences or individual job search, Figure IXb replicates IXa for workers with deductions between DKr 7,500 and DKr 25,000. Note that these workers' tax incentives change at neither the statutory kink nor the pension kink. These workers exhibit no excess propensity to locate near the pension kink ($b = -0.01$), implying that there is little aggregate bunching at the pension kink. In contrast, Figure IXc shows that the *same* workers exhibit substantial bunching around the statutory kink ($b = 0.56$), confirming that there is significant aggregate bunching at the statutory kink. Together, these figures offer two lessons. First, the bunching at the pension kink is driven by individual job search – i.e., finding a job that pays DKr 33,000 above the top kink – rather than distortions in the distribution of offers. Second, aggregate bunching is significant only at kinks that affect large groups of workers, consistent with the model's prediction that the distribution of job offers is tailored to match aggregate worker preferences.

One of the reasons that 60% of individuals face the statutory top tax kink is that the top tax is based on individual earnings. The scope of the middle tax cutoff is smaller because it depends upon household income; 38% of individuals in the economy begin to pay the middle tax when their income crosses the statutory middle tax cutoff. This raises the concern that there may be less bunching at the middle kink than the top kink not just because it has smaller size but also because it has smaller scope. To distinguish size from scope, we compare bunching at the middle tax pension kink (the point at which individuals who are at the pension cap begin paying the middle tax) with bunching at the top tax pension kink. Both of these kinks affect very few workers in the economy (i.e. have scope near zero), but the top tax pension kink is much larger in size than the middle tax pension kink. We find that there is no bunching ($b = -0.01$) in wage earnings at the middle tax pension

30. We condition on having deductions greater than DKr 20,000 to isolate the relevant part of the population in order to detect bunching at the pension kink. To allay the concern that conditioning on deductions greater than DKr 20,000 creates selection bias, we verified that conditioning on deductions in the previous year produces similar results ($b = 0.54$). We also ran a series of placebo tests conditioning on having deductions above thresholds ranging from −20,000 to 40,000 and found no bunching at any points in the wage earnings distribution except for the statutory kink and the pension kink.

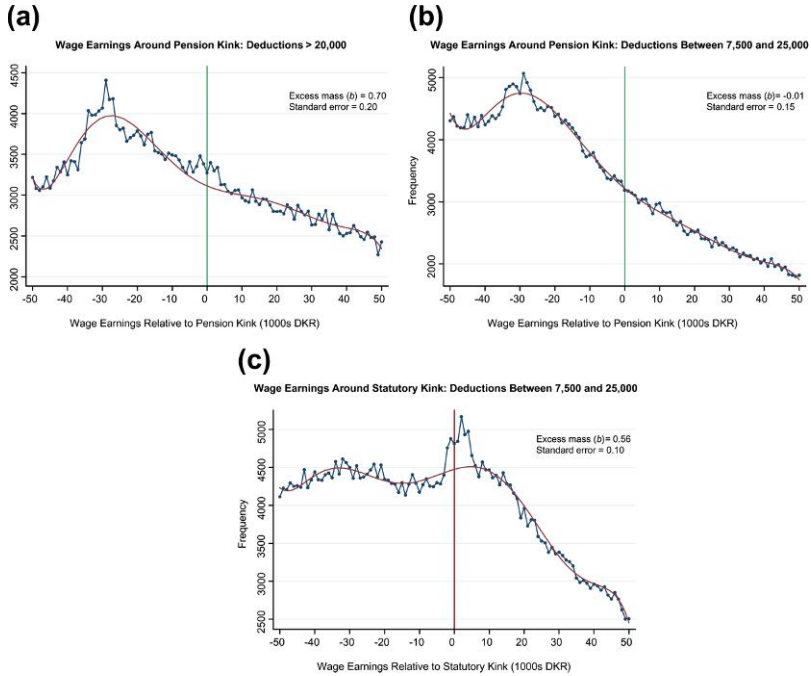

FIGURE IX
Individual vs. Aggregate Bunching at the Pension Kink

Panel (a) plots the distribution of wage earnings relative to the pension kink (demarcated by the light gray vertical line) for wage earners with greater than DKr 20,000 of net deductions. The pension kink is defined as the top tax bracket cutoff plus the maximum tax-deductible pension contribution in each year. Panel (b) replicates (a) for wage earners with between DKr 7,500 and DKr 25,000 of net deductions. Panel (c) plots the distribution of wage earnings relative to the statutory top kink (demarcated by the dark gray vertical line) for wage earners with between DKr 7,500 and DKr 25,000 in net deductions. The figure also shows the counterfactual distributions and excess masses, computed as in Figure IIIa.

kink (see Figure A.7), supporting prediction 1 by showing that size matters holding scope fixed.[31]

We now turn to the third part of prediction 2: do workers with small deductions bunch more than those with large deductions? The econometric challenge in testing this prediction is that

31. The lack of individual bunching at the middle tax pension kink also explains why there is no aggregate bunching at the middle tax kink: firms have no reason to offer jobs at the kink if workers themselves do not demand such jobs. Firm responses amplify bunching only if the kink is large enough to induce individual bunching to begin with.

deductions themselves are endogenous. In particular, workers with large deductions may have chosen their deductions in order to reach the top tax kink. We address this endogeneity problem using a grouping instrument. We compute the fraction of workers with deductions less than DKr 7,500 in magnitude for cells of the population defined by marital status, gender, year, and age (in decades). We then divide workers into ten equal-width bins based on the fraction of workers with small deductions in their group and estimate the degree of bunching at the top kink (b) for workers in each of these ten bins.[32] Figure X plots the estimated b vs. the fraction of workers with small deductions in the ten groups. The groups with small deductions exhibit much greater bunching: the slope of the fitted line in Figure X is statistically significant with $p < 0.01$. This result confirms that tax incentives that affect a larger group of workers generate large observed elasticities. Workers with small deductions can rely on aggregate bunching to reach the top kink, whereas workers with large deductions need to actively search for a less common job.

Further supporting the importance of aggregate bunching, we find that some of the heterogeneity in elasticities across demographic groups (as in Figure IIIb) is driven by occupational choice. For instance, reweighting men's occupations to match those of women's eliminates 50% of the gap in observed elasticities between men and women (see Online Appendix Figure A.8).

Changes in the aggregate distribution of job offers also shape earnings dynamics as the tax bracket changes. To characterizedefine an indicator for whether an individual's change in wage earnings from year t to year $t + 2$ is within DKr 7500 (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 tracks the movement in the kink over time. Figure XIa plots the fraction of individuals who track the movement in the kink vs. the level of wage earnings in the base year relative to the statutory kink. The propensity to track the movement in the kink is highest for individuals near the kink to begin with. Figure XIb replicates Figure XIa for the pension kink, focusing on individuals with deductions greater than 20,000 in year t, as in Figure IXa. Individuals at the pension kink in year t do not have any excess

32. We exclude groups with a fraction of workers with small deductions in the bottom and top 5% of the distribution, as there are too few observations to estimate b in equal-width bins in the tails.

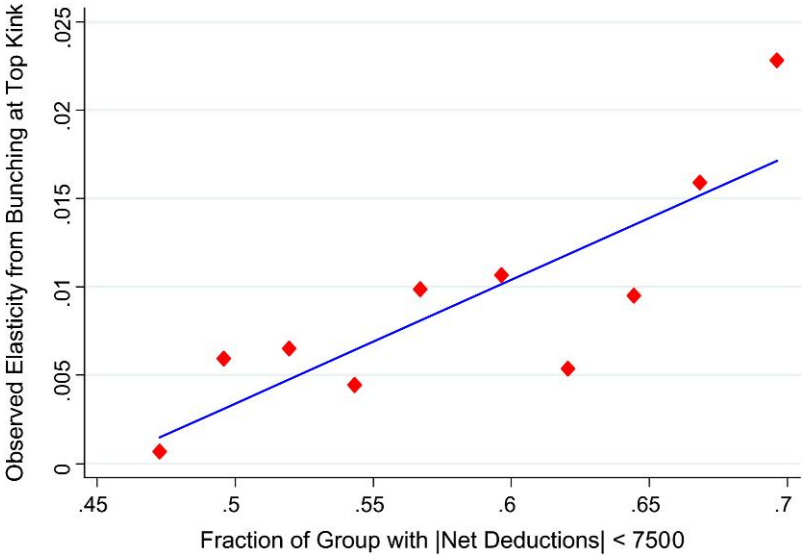

FIGURE X
Observed Elasticities vs. Scope of Tax Changes

To construct this figure, we first calculate the fraction of individuals with net deductions less than DKr 7,500 in magnitude in each age-gender-marital status-year cell. We then group individuals into 10 equal-width bins based on the fraction with small deductions in their group as described in the text. We estimate the excess mass at the top kink as in Figure IIIa and apply equation (6) to calculate observed elasticities for each of the ten groups. The figure shows a scatter plot of the observed elasticities vs. the fraction with small deductions in the 10 bins. The best-fit line is estimated using OLS.

propensity to track the movement in the pension kink. Instead, aggregate bunchers at the statutory kink (located at approximately DKr $-33,000$ in Figure XIb), exhibit a higher propensity to move with the kink even though they have no incentive to do so. In sum, individuals who reach the kink via aggregate bunching move with the kink whereas those who get there through individual job search do not. Intuitively, firms adjust the packages they offer as the aggregate distribution of workers' tax preferences change, whereas workers must pay search costs to switch jobs and actively track the kink themselves.[33]

33. These results also provide further evidence that the difference in bunching at the top and middle kinks is not driven by heterogeneous elasticities. If individuals near the top tax cutoff were simply more elastic and did not face adjustment costs, they would track the movement of the top kink over time.

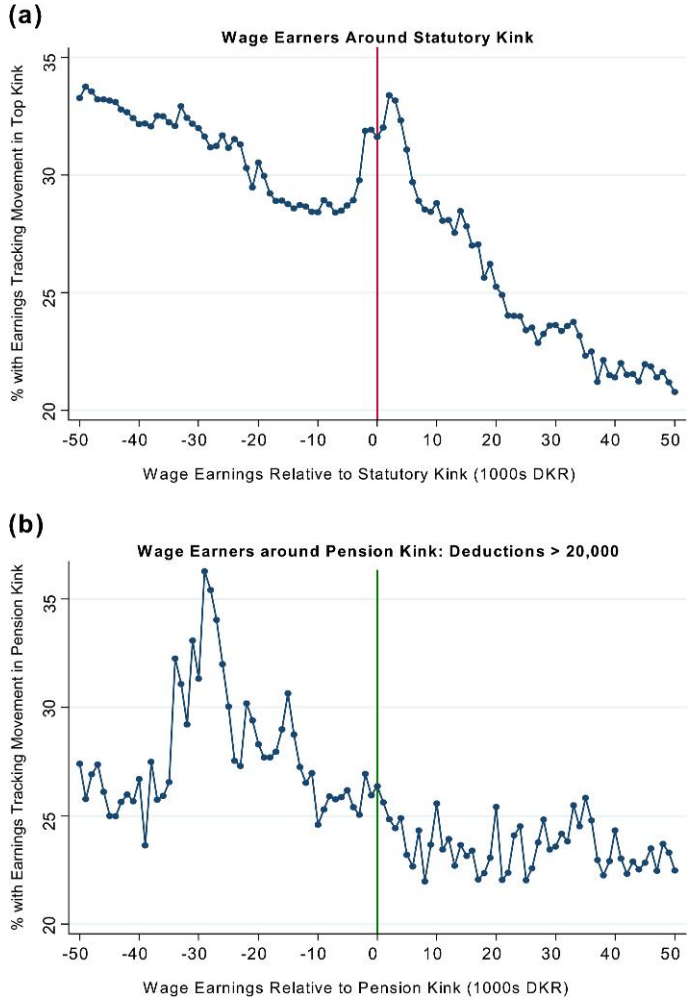

FIGURE XI
Dynamics of Earnings Around the Top Tax Cutoff

These figures show how the propensity to track the movement in the top tax cutoff across years varies across individuals. To construct Panel (a), we first divide individuals into bins of DKr 1000 in wage earnings in a given year t, and calculate the fraction in each bin whose change in wage earnings from a year t to $t + 2$ falls within DKr 7,500 of the movement in the top tax bracket cutoff from year t to $t + 2$. Panel (a) plots this fraction for wage earnings bins around the statutory top tax cutoff. Panel (b) replicates (a) for the pension kink, restricting the sample to wage earners with net deductions greater than DKr 20,000. It shows the fraction of individuals whose change in wage earnings falls within DKr 7,500 of the movement in the pension kink for wage earnings bins around the pension kink.

We conclude that firm responses play a central role in shaping the effects of tax changes on equilibrium labor supply. Such responses may be particularly easy to detect in Denmark because collective bargaining facilitates the aggregation of workers' tax preferences. While collective bargaining is less common in economies such as the U.S., technological constraints lead to hours constraints in all labor markets. The general lesson to be drawn from the evidence here is that these constraints are endogenous to the tax regime.

IV.C. Prediction 3: Correlation Between Individual and Aggregate Bunching

We test the third prediction of the model by examining the correlation between individual and aggregate bunching across occupations. As above, we measure aggregate bunching b_q^A in occupation q by measuring the excess mass in the wage earnings distribution at the statutory top tax cutoff for individuals who have more than DKr 20,000 in deductions (and therefore have no incentive to locate at the statutory kink). We measure individual bunching b_q^I by the excess mass at the pension kink in the wage earnings distribution for individuals in occupation q with more than DKr 20,000 in deductions, because this kink has near-zero scope ($\zeta \simeq 0$). Note that b_q^A and b_q^I are estimates of bunching at two *different* kinks for the same group of individuals, and thus are not mechanically related.

Figure XII plots the estimates of b_q^A vs. estimates of b_q^I across occupations defined at the 2 digit ISCO level. The (unweighted) correlation between b_q^A and b_q^I is 0.65 and is significantly different from 0 with $p < 0.001$. In a regression weighted by occupation size, 64% of the variation in b_q^A is explained by the variation in b_q^I. Note that the few negative point estimates of b_q^I and b_q^A are not significantly different from zero. We cannot interpret the positive correlation in Figure XII as evidence that differences in individuals' preferences *cause* changes in the distribution of jobs offered as they could also be driven by sorting of workers into occupations that suit their tastes. Nevertheless, the evidence is consistent with the model's prediction that firms (or unions) cater to workers' tax-distorted preferences in equilibrium.

IV.D. Self-Employed Individuals

The self-employed are a useful comparison group because they face much smaller frictions in adjusting taxable income than wage

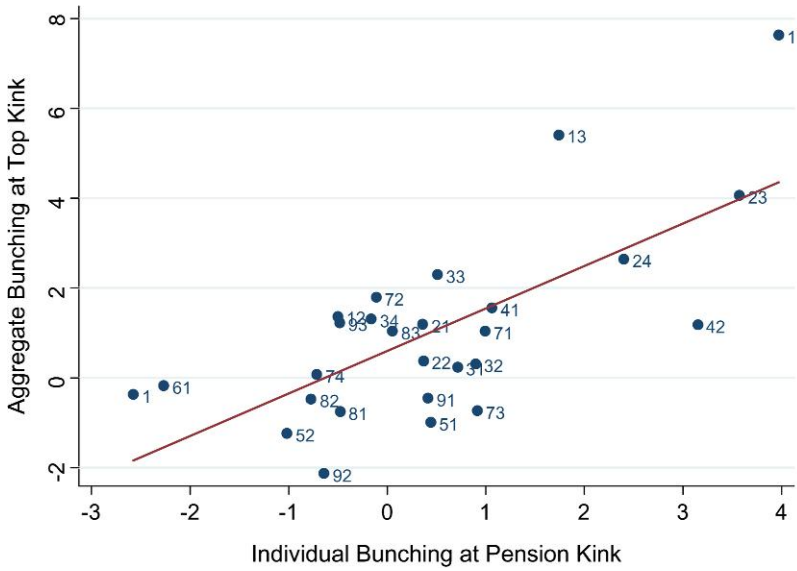

FIGURE XII
Correlation Between Individual and Aggregate Bunching

This figure plots the amount of aggregate bunching (b_q^A) vs. the amount of individual bunching (b_q^I) for all International Standard Classification of Occupation codes at the two digit level. Both aggregate and individual bunching are estimated on the subgroup of individuals with net deductions greater than DKr 20,000, as in Figure IXa. Individual bunching is the excess mass at the pension kink for this group, while aggregate bunching is the excess mass at the statutory top tax cutoff for the same group. See Table A.1 for a list of the occupation codes.

earners. They are not subject to hours constraints and do not need to search for a different job to change their earnings. They can also easily change reported taxable incomes, either by shifting realized income across years or by under-reporting taxable incomes.[34] Therefore, we expect that the model's three predictions should not apply to the self-employed.

Figure XIII replicates the key graphs shown above for the self-employed. Figure XIIIa shows that the self-employed exhibit extremely sharp bunching at the top kink, consistent with their ability to adjust their income more easily. The estimated excess mass is b = 18.4 at the top kink, dwarfing the excess mass

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

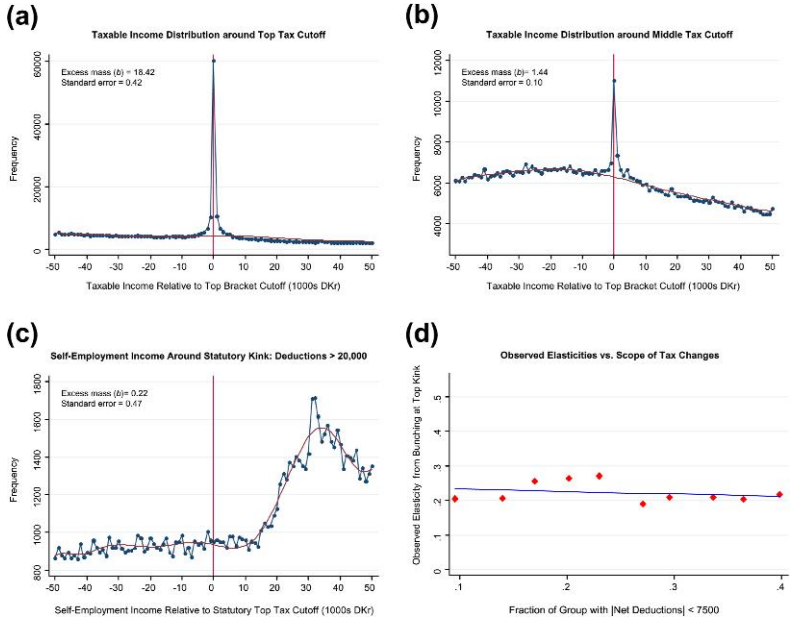

FIGURE XIII
Self-Employed Individuals

These figures include only individuals who report positive self-employment income. Panels (a) and (b) plot the taxable income distribution around the top and middle cutoffs from 1994–2001. Panel (c) plots the distribution of realized self-employment income around the statutory top tax cutoff for individuals with net deductions greater than 20,000. Panel (d) replicates Figure X for individuals with positive self-employment income, with the y axis scaled to have the same range relative to the mean observed elasticity as in Figure X.

for wage earners and implying an observed elasticity of 0.24. Figure XIIIb shows that unlike wage earners, the self-employed also bunch sharply at the middle tax kink. The observed elasticity at the middle kink is 0.10. We believe that the observed elasticity at the middle kink is smaller than that at the top kink because capital income is subject to the middle tax but not the top tax. Self-employed individuals are allowed to reclassify some of their profits as capital income, creating an added margin of response at the top tax cutoff. Consistent with this explanation, self-employed individuals with capital income less than DKr 1,000 in magnitude have an observed elasticity of 0.16 at the middle kink vs. 0.20 at the top kink.

Figure XIIIc tests for aggregate bunching by plotting the income distribution around the statutory kink for self-employed individuals with deductions larger than DKr 20,000. Unlike wage earners, self employed individuals with large deductions exhibit no excess mass around the statutory kink. As a result, self employed individuals with common tax preferences (small deductions) bunch just as much as those with uncommon tax preferences (large deductions). This is shown in Figure XIIId, which is constructed using mean group deductions in the same way as Figure X.

These "placebo tests" confirm that our three predictions do not apply to the self-employed.[35] Some of the bunching among the self-employed is driven by intertemporal shifting and evasion. LeMaire and Schjerning (2007) demonstrate using the same Danish data that the self-employed adjust their retained earnings and profit distributions over time to remain below the top tax threshold in each year. Kleven et al. (2010) uncover substantial tax evasion among the self-employed and estimate that 40% of the bunching at the top kink is driven by tax evasion. Eliminating this evasion component of bunching at the top kink implies a taxable income elasticity for the self employed of 0.14. Regardless of which margin the self employed use, we can conclude that frictions significantly attenuate observed elasticities: the size and scope of tax changes matters less for margins of behavior with low frictions (changing reported taxable income or self-employment earnings) than for margins with higher frictions (changing wage earnings).

V. Conclusion

This paper has shown that the effects of tax policies on labor supply are shaped by adjustment costs and hours constraints endogenously chosen by firms. Because of these forces, modern microeconometric methods of estimating elasticities – focusing on policy changes that affect a subgroup of workers – may underestimate the "structural" elasticities that control steady-state responses.

35. Furthermore, we find that individuals who switch between self-employment and wage earning have a much greater propensity to bunch at kinks in the years when they are self employed.

Our empirical analysis does not yield an estimate of the structural (macro) elasticity. In Chetty et al. (2009), we calibrate a more general version of the model presented here. We find that the structural elasticity that matches the evidence is an order of magnitude larger than the observed elasticity at the top kink. Intuitively, a small ε cannot produce substantial variation in observed elasticities across tax changes of different size and scope because the costs of deviating from optimal hours are very large when ε is small. In future work, it would be useful to identify ε more precisely by structurally estimating a more realistic dynamic model of labor supply with frictions.

It would also be interesting to explore the normative implications of adjustment costs and firm responses. For example, the efficiency cost of a tax levied on one group of workers may depend not just upon their elasticities but also upon those of their co-workers if firms are constrained to offer similar packages to different workers. 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, Ichino and Karabarbounis 2007; Kleven, Kreiner, and Saez 2009). If the difference in observed elasticities across genders is caused by heterogeneity in occupational frictions rather than tastes, there may be less justification for higher taxes on secondary earners in steady state.

Finally, the results here call for caution in using quasi-experiments that apply to small subgroups to learn about the effects of economic policies on behavior. In settings with rigid institutional structures and frictions in adjustment, the steady-state effects of policies implemented at an economy-wide level could differ substantially from the effects of such experiments.

HARVARD UNIVERSITY AND NBER
HARVARD UNIVERSITY AND NBER
HARVARD UNIVERSITY AND CAM
STANFORD UNIVERSITY AND NBER

REFERENCES

Altonji, Joseph G., and Christina H. Paxson, "Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-Offs," *Journal of Labor Economics*, 6 (1988), 254–276.

Alesina, Alberto F., Edward L. Glaeser, and Bruce Sacerdote, "Work and Leisure in the U.S. and Europe: Why So Different?" in *NBER Macroeconomics Annual*, eds. M. Gertler and K. Rogoff, 20 (2005), 1–64.
Alesina, Alberto F., Andrea Ichino, and Loukas Karabarbounis, "Gender-Based Taxation and the Division of Family Chores," NBER Working Paper 13638 (2007).
Andersen, Torben M., and Michael Svarer, "Flexicurity: Labour Market Performance in Denmark," *CESifo Economic Studies,* 53 (2007), 389–429.
Blundell, Richard, and Thomas MaCurdy, "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics*, Orley Ashenfelter and David Card eds., (1999).
Boskin, Michael J., and Eytan Sheshinski, "Optimal Tax Treatment of the Family: Married Couples," *Journal of Public Economics,* 20 (1983), 281–297.
Chetty, Raj, "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," NBER Working Paper No. 15616 (2011).
Chetty, Raj, "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance," *American Economic Journal: Economic Policy,* 1 (2009), 31–52.
Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," NBER Working Paper No. 15617, 2009.
Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber, "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins." *American Economic Review,* forthcoming, May 2011.
Chetty, Raj, and Emmanuel Saez, "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients," NBER Working Paper No. 14836, 2009.
Cogan, John, "Fixed Costs and Labor Supply," *Econometrica*, 49 (1981), 945–963.
Davis, Steven, and Magnus Henrekson, "Tax Effects on Work Activity, Industry Mix and Shadow Economy Size: Evidence from Rich-Country Comparisons," in *Labour Supply and Incentives to Work in Europe*, Ramon Gomez Salvador, Ana Lamo, Barbara Petrongolo, Melanie Ward, and Etienne Wasmer eds. (Northamption, MA: Edward Elgar Publishing, 2005).
Dickens, William T., and Shelly J. Lundberg, "Hours Restrictions and Labor Supply," *International Economic Review*, 34 (1993), 169–192.
Earle, John S. and John Pencavel, "Hours of Work and Trade Unionism," *Journal of Labor Economics* 8 (1990), S150–74.
Farber, Henry S., "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers," *Journal of Political Economy*, 113 (2005), 46–82.
Fehr, Ernst, and Lorenz Goette, "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment," *American Economic Review,* 97 (2007), 298–317.
Feldstein, Martin, "Tax Avoidance and the Deadweight Loss of the Income Tax." *Review of Economics and Statistics*, 81 (1999), 674–680.
Goolsbee, Austin, "What Happens When You Tax the Rich? Evidence from Executive Compensation," *Journal of Political Economy*, 108 (2000), 352–378.
Gruber, Jonathan, and Emmanuel Saez, "The Elasticity of Taxable Income: Evidence and Implications," *Journal of Public Economics*, 84 (2002), 1–32.
Ham, John C., "Estimation of a Labour Supply Model with Censoring Due to Unemployment and Underemployment," *Review of Economic Studies*, 49 (1982), 335–354.
Hansen, Claus Thustrup. "Lower Tax Progression, Longer Hours and Higher Wages." *Scandinavian Journal of Economics*, 101 (1999), 49–65.
Hausman, Jerry A., "Labor Supply," in *How Taxes Affect Economic Behavior*, Henry Aaron and Joseph Pechman eds. (Washington D.C.: The Brookings Institution, 1981).
Holmlund, Bertil, and Martin Söderström, "Estimating Dynamic Income Responses to Tax Reforms: Swedish Evidence," mimeo, Uppsala University, 2008.
Kleven, Henrik, and Claus Kreiner, "Welfare and Taxation" in *Tax, Work, and Equality: An Investigation of the Danish Tax and Welfare System,* Torben Tranaes ed. (Copenhagen: Gyldendal, 2006)

Kleven, Henrik, Martin Knudsen, Claus T. Kreiner, Søren Pedersen, and Emmanuel Saez, "Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark," NBER Working Paper 15769, 2010.
Kleven, Henrik, Claus Kreiner, and Emmanuel Saez, "The Optimal Income Taxation of Couples," *Econometrica*, 77 (2009), 537–560.
Kleven, Henrik and Esben A. Schultz, "Estimating Taxable Income Responses Using Danish Tax Reforms," mimeo, London School of Economics, 2010.
LeMaire, Daniel, and Bertel Schjerning, "Earnings, Uncertainty, and the Self-Employment Choice," CBER Discussion Paper DP 2007–04, 2007.
Lindbeck, Assar, "Welfare State Disincentives with Endogenous Habits and Norms," *The Scandinavian Journal of Economics*, 97 (1995), 477–494.
MaCurdy, Thomas, David Green, and Harry Paarsch, "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply," *The Journal of Human Resources*, 25 (1990), 415–490.
Ministry of Taxation, *Skatteberegningsreglerne, før og nu*, http://www.skm.dk/public/dokumenter/Tal45/Skatteberegning90/Skatteberegningsreglerne_2.pdf, 2002.
Moffitt, Robert A., "The Econometrics of Kinked Budget Constraints," *Journal of Economic Perspectives*, 4 (1990), 119–139.
Oettinger, Gerald S., "An Empirical Analysis of the Daily Labor Supply of Stadium Vendors," *Journal of Political Economy*, 107 (1999), 360–392.
Rogerson, Richard. "Comment on Work and Leisure in the U.S. and Europe: Why so different?", in *NBER Macroeconomics Annual*, eds. M. Gertler and K. Rogoff, 20 (2005), 79–95.
Rogerson, Richard, and Johanna Wallenius, "Micro and Macro Elasticities in a Life Cycle Model with Taxes," *Journal of Economic Theory,* 144 (2009), 2277–2292.
Rosen, Harvey S., "Taxes in a Labor Supply Model with Joint Wage-Hours Determination," *Econometrica* , 44 (1976), 485–507.
Saez, Emmanuel, "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy,* 2 (2010), 180–212.
Saez, Emmanuel, Joel B. Slemrod, and Seth H. Giertz, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature,* forthcoming, 2010.
Slemrod, Joel B., "Income Creation or Income Shifting? Behavioral Responses to the Tax Reform Act of 1986," *American Economic Review,* 85 (1995), 175–180
Slemrod, Joel B., and Shlomo Yitzhaki, "Tax avoidance, evasion, and administration," in *Handbook of Public Economics*, Alan Auerbach and Martin Feldstein eds. (Amsterdam: North-Holland, 2002).