NBER WORKING PAPER SERIES

MICRO VS MACRO LABOR SUPPLY ELASTICITIES:
THE ROLE OF DYNAMIC RETURNS TO EFFORT

Henrik Kleven
Claus Thustrup Kreiner
Kristian Larsen
Jakob Egholt Søgaard

Working Paper 31549
http://www.nber.org/papers/w31549

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2023, Revised March 2025

We thank Raj Chetty, Jeppe Druedahl, John Friedman, Peter Ganong, Nathan Hendren, Pat Kline, Wojciech Kopczuk, Ilyana Kuziemko, Camille Landais, Felipe Lobel, Benjamin Lockwood, Emi Nakamura, Richard Rogerson, Benjamin Schoefer, Johannes Spinnewijn, Damián Vergara, Danny Yagan, Owen Zidar, and two anonymous referees for comments. The activities of CEBI are financed by the Danish National Research Foundation, grant DNRF134. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Micro vs Macro Labor Supply Elasticities: The Role of Dynamic Returns to Effort
Henrik Kleven, Claus Thustrup Kreiner, Kristian Larsen, and Jakob Egholt Søgaard
NBER Working Paper No. 31549
August 2023, Revised March 2025
JEL No. C1, D6, E6, H2, H3, J2, J3

ABSTRACT

We investigate long-run earnings responses to taxes in the presence of dynamic returns to effort. First, we develop a theoretical model of earnings determination with dynamic returns to effort. In this model, earnings responses are delayed and mediated by job switches. Second, using administrative data from Denmark, we verify our model's predictions about earnings and hours-worked patterns over the lifecycle. Third, we provide a quasi-experimental analysis of long-run earnings elasticities. Informed by our model, the empirical strategy exploits variation among job switchers. We find that the long-run elasticity is around 0.5, considerably larger than the short-run elasticity of roughly 0.2.

Henrik Kleven
Department of Economics
Princeton University
238 Julis Romo Rabinowitz Building
Princeton, NJ 08544
and CEPR
and also NBER
kleven@princeton.edu

Claus Thustrup Kreiner
Department of Economics
University of Copenhagen and CEBI
Øster Farimagsgade 5
Building 35
DK-1353 Copenhagen K
Denmark
ctk@econ.ku.dk

Kristian Larsen
University of Copenhagen and CEBI
Department of Economics
Øster Farimagsgade 5
Building 35
DK-1353 Copenhagen K
Denmark
kuol@econ.ku.dk

Jakob Egholt Søgaard
University of Copenhagen and CEBI
Øster Farimagsgade 5
Building 26
DK-1353 Copenhagen K
Denmark
jes@econ.ku.dk

1 Introduction

The impact of taxes on labor supply and earnings is critical for assessing the equality-efficiency trade-off and optimal redistribution. The strength of these responses is typically measured by the elasticity of earnings with respect to taxes. Importantly, the welfare-relevant elasticity captures long-run responses, accounting for tax-induced changes in the full lifetime trajectory of earnings. Such responses are challenging to estimate and there is no consensus on plausible magnitudes. Even if much microeconometric evidence points to small elasticities, many economists contend that the true long-run effect of taxes is large due to dynamic effects (e.g., Prescott 2004).

This debate relates to the stark difference between micro and macro elasticities of labor supply. Micro estimates — typically using tax reforms as quasi-experiments — tend to be small. Macro estimates on the other hand — typically based on structural estimations or calibrations — tend to be large. To illustrate the extent of disagreement between these two research strands, consider the Laffer rate on top earners. The top-income Laffer rate in the US is close to 80% if the elasticity is 0.2 (a typical micro estimate), but only 40% if the elasticity is 1 (a typical macro estimate).[1] This range is too large to provide useful policy guidance.

Which approach is right and which is wrong? Our starting point is that both are right and both are wrong. Micro studies are based on research designs that allow for causal identification, but the approach only captures short-run effects and may miss important dynamic mechanisms. Macro studies are model-dependent and may be associated with specification bias, but they allow for potentially relevant dynamic responses. The goal of this paper is to develop a quasi-experimental approach that is better able to capture welfare-relevant, long-run elasticities.

We are particularly interested in the elasticity at the top of the income distribution, among salaried career workers. Such top earners represent a large fraction of income and tax revenue, making them critical to tax design. A key challenge to estimating their responsiveness to taxes is the presence of dynamic returns to effort. Consider an example close to home: top academics. An

[1] These numbers are based on the Laffer rate formula $\tau = 1/(1+\varepsilon a)$, where ε is the earnings elasticity and a is the Pareto parameter (Diamond 1998; Saez 2001). The Pareto parameter is about 1.5 in the US. An elasticity of 1 is, if anything, a conservative assessment of the macro literature (see e.g., Prescott 2004; Rogerson and Wallenius 2009; Keane and Rogerson 2015). Instead of using the Laffer rate to illustrate the policy implications of elasticity uncertainty, we could consider the implications for the Marginal Cost/Value of Public Funds (MCPF/MVPF) as analyzed by Kleven and Kreiner (2006) and Hendren and Sprung-Keyser (2020). The difference between micro and macro elasticities will in general imply widely different MCPFs.

academic career consists of doing research and building a publication record, which may eventually lead to promotions within a department or outside offers from other departments. The link between effort and earnings is delayed and discrete, centered around promotions or firm switches. We posit that such dynamic and discrete returns characterize most top professions. Standard quasi-experimental research designs are largely uninformative in the presence of dynamic returns. They rely on models where outcomes respond almost immediately to incentives such as in the static and frictionless model of hourly-paid workers.

Our agenda relates to an idea previously studied in macro and structural labor economics: the effect of effort on human capital accumulation via learning-by-doing and on-the-job training. Structural estimations of human capital models are consistent with large long-run elasticities (see e.g., Keane 2011; Keane and Rogerson 2015).[2] While human capital accumulation is one way of generating dynamic returns, a variety of other mechanisms may be at play. In the example of top academics, it is not *a priori* clear if wages increase over the career path due to changes in productivity or because discrete performance evaluations reward historical output. We take the latter view, the implication of which is that the lifecycle profile of earnings is a step function with discrete changes at job switches such as occupation or firm switches. This idea is related to a literature in labor economics showing that job-to-job mobility is central for earnings growth over the lifecycle either through gains in the job-match component of wages (e.g., Topel and Ward 1992; Pavan 2011) or through mobility to firms with higher wage premia (e.g., Abowd, Kramarz, and Margolis 1999; Card, Heining, and Kline 2013; Card, Cardoso, and Kline 2016; Card, Cardoso, Heining, and Kline 2018).[3] We develop a new model that highlights the role of dynamic returns realized at the point of switching and investigate the implications for the estimation of behavioral responses to taxes.

Our model distinguishes between *realized* earnings and *latent* earnings (effort). Workers make effort choices based on their productivity and taxes, taking into account that higher effort generates higher earnings with a delay. Realized earnings change only at discrete job events — such as switches between occupations or firms — at which time realized and latent earnings are realigned. We consider a benchmark model where the probability of switching is exogenous and an extension where this probability is endogenized. The standard labor supply model obtains as a special case

[2]Best and Kleven (2013) develop a theory of optimal taxation in a setting where effort affects future wages through human capital accumulation.

[3]A macro literature also provides evidence on the importance of job switches for wage and earnings growth (e.g., Grigsby, Hurst, and Yildirmaz 2021; Guvenen, Karahan, Ozkan, and Song 2021).

with a switching probability of one, in which case the elasticity of true effort η (which governs the long-run macro elasticity) is identical to the elasticity of realized earnings ε (the short-run micro elasticity). Outside this limit case, the macro elasticity is larger than the micro elasticity. Allowing for heterogeneity in both structural elasticities η and switching probabilities λ, we characterize the conditions under which the macro elasticity can be point identified or partially identified using responses among short-run switchers. The macro elasticity can be point identified when η and λ are orthogonal and partially identified when η and λ are correlated. The two cases can be separated based on the relationship between observed micro elasticities and the timing of switches following a tax reform.

The empirical part of the paper leverages Danish administrative data to verify the predictions of the model and identify the long-run macro elasticity. The data are employer-employee matched and contain detailed occupation codes, allowing us to observe job switches at a granular level. We start by providing descriptive evidence on earnings and hours-worked patterns over the lifecycle, highlighting three facts of the data. First, earnings are strongly related to past hours worked, conditional on current hours worked and rich controls for demographics and skill. Considering workers at advanced career stages, past hours is a stronger predictor of earnings than current hours. Second, for workers who reach the top of the distribution, the lifecycle profile of earnings is discrete, driven by jumps at job switches and inaction between switches. In fact, between-job variation accounts for almost 95% of the total variation in earnings over the lifecycle. Finally, based on event studies of promotions — defined as switches to job cells with higher median earnings — we show that individual earnings jump discretely at promotion events while hours worked are smooth.[4] These facts are consistent with our theoretical model, but not with standard models.

Informed by our model and descriptive evidence, we provide a quasi-experimental study of earnings elasticities using firm×occupation switchers. The main analysis is based on a large tax reform implemented in Denmark in 2009. This reform reduced the top marginal tax rate by about 11pp, while leaving tax rates further down the distribution roughly unchanged.[5] We estimate behavioral responses using difference-in-differences and triple-differences approaches. Considering the full population of treated workers, we find clear and precisely estimated earnings responses to taxes. But the average response is modest in elasticity terms. We then split the data into job

[4]Related, Bronson and Thoursie (2021) show that discrete earnings jumps within firms account for a large fraction of the gender gap in earnings.

[5]To address concerns that the reform was passed at the time of the great recession, we consider a historical tax reform — the 1987-reform — which also reduced marginal tax rates at the top relative to the bottom. The estimates from the 1987- and 2009-reforms are virtually identical.

movers and job stayers, showing that the small average response masks striking heterogeneity: job movers feature large responses — an elasticity of about 0.5 — while job stayers feature no responses. Because a minority of workers switch in any given year, the large switcher elasticity is consistent with a small average elasticity.

Our central thesis is that earnings responses among short-run switchers can be used to uncover the long-run elasticity in the population. As mentioned, point identification requires orthogonality between structural elasticities η and switching probabilities λ, the implication of which is that the observed switcher elasticity ε is constant in the timing of switching. We estimate impacts by the timing of job switching following the reform, showing that the time profile of the switcher elasticity is virtually flat. This is consistent with point identification.

We provide three additional analyses to address possible selection bias. First, we estimate the effect of the tax reform on the probability of switching, finding small effects on this margin. This allows us to bound the selection bias that would arise if these reform-induced switchers are selected on the structural elasticity η. We show that, even under extreme assumptions about such selection, it has only small effects on our estimates. The simple reason is that the marginal (and potentially selected) switchers constitute a very small fraction of all switchers. Second, even though the reform had a small effect on the amount of switching, it could have changed the composition of switchers. To investigate this threat, we estimate impacts of the reform on the characteristics of job switchers. Looking at a wide range of variables, we do not find effects on any of them. Finally, we restrict the sample to plausibly exogenous switches, namely those triggered by mass layoffs. Mass-layoff switchers feature similar earnings responses to the tax reform as the full sample of switchers.[6] Taken together, this set of findings strongly suggests that our estimates are not biased by selection.

Our paper contributes to several literatures in public finance, labor, and macroeconomics. First, we contribute to an enormous body of work estimating labor supply elasticities with respect to tax incentives, as reviewed by Blundell and MaCurdy (1999) and Saez, Slemrod, and Giertz (2012). Summarizing the microeconometric evidence, Saez, Slemrod, and Giertz (2012) argue that "the profession has settled on a value for this elasticity close to zero."[7] Consistent with this view, we

[6] Importantly, the positive earnings response to lower taxes for mass-layoff switchers is based on our quasi-experimental design, which compares treated and untreated mass-layoff switchers from before to after the reform. As we show, this earnings response is consistent with a negative reduced-form effect of a mass layoff itself (see e.g., Jacobson, LaLonde, and Sullivan 1993).

[7] This assessment pertains to elasticities of *real* labor supply (or wage earnings) along the *intensive margin*, consistent with the focus of our study. Estimates of taxable income elasticities — including avoidance and evasion responses — can be considerably larger depending on the tax code and enforcement system. Estimates of extensive margin elasticities

estimate a small earnings elasticity when taking a conventional quasi-experimental approach.[8] We argue that such micro elasticities are uninformative of long-run responses among top earners, most of whom work in salaried jobs. Such workers cannot freely adjust earnings within a given job cell. They may change effort, but the earnings implications of changed effort play out dynamically and are often tied to job switches. Using hours worked as the outcome variable is not a solution because, for salaried workers, hours is a very limited measure of true effort.[9] Rather, our proposed solution is to restrict attention to job switchers, maintaining earnings as the outcome variable. While some papers have studied heterogeneity in tax elasticities by job switching status (e.g., Tortarolo, Cruces, and Castillo 2020), we are not aware of any work that develops a theoretical framework and empirical approach using switchers to estimate welfare-relevant, long-run earnings elasticities.[10]

Our paper presents a new attempt to reconcile micro and macro evidence on labor supply. The micro-macro debate has focused on three issues: extensive margin responses (Chetty, Guren, Manoli, and Weber 2013), optimization frictions (Chetty 2012), and human capital accumulation (Imai and Keane 2004; Keane 2011; Keane and Rogerson 2015). Our approach is related to models incorporating human capital effects of effort — a specific channel through which dynamic returns may arise — but is at the same time fundamentally different. In standard human capital models, worker compensation is aligned with actual effort and productivity at any point in time, where productivity is allowed to change over time due to learning-by-doing or on-the-job training. Such effects are presumably slow-moving, and there is no role for discrete changes in earnings around job switches. Our approach using short-run switchers is not plausibly driven by human capital effects, while the human capital literature does not capture the effects studied here. Although we argue that the long-run elasticity is larger than typical micro estimates, our estimates remain considerably smaller than those implied by a number of macro studies.[11]

feature much greater variation across studies and less of a consensus (see Kleven 2024).

[8] Studying the same Danish tax reform, Kreiner, Leth-Petersen, and Skov (2016), Jakobsen and Søgaard (2022), Labanca and Pozzoli (2022), and Sigaard (2023) also estimate small micro elasticities.

[9] In fact, this is one of the main reasons why the modern public finance literature has shifted its focus from hours-of-work elasticities to earnings elasticities. But by doing so, researchers solved one problem (the fact that hours responses are too narrow) by introducing another one (the fact that earnings responses are dynamic and delayed).

[10] A key advantage of our approach is that it relies on a widely available source of tax variation. This allows for applying the approach across different countries, reforms, and samples. A different approach would be to hunt for the "perfect" experiment to estimate long-run earnings responses. A compelling example of such an approach is provided by Saez, Matsaganis, and Tsakloglou (2012). They consider a cohort-based payroll tax reform in Greece and find zero long-run earnings responses. We are motivated by the view that such experiments are rare and context-specific. Building a concensus regarding the magnitude of long-run elasticities requires an approach that can be implemented repeatedly across settings and over time.

[11] Our agenda is also complementary to a paper by Scheuer and Werning (2017) on the optimal taxation of superstars.

Our work is also related to the literature studying how optimization frictions shape observed labor supply. This includes a labor literature on hours constraints and adjustment costs (Altonji and Paxson 1986, 1988; Lachowska, Mas, Saggio, and Woodbury 2023) and a public finance literature showing that micro elasticities may be strongly attenuated by frictions (Chetty, Friedman, Olsen, and Pistaferri 2011; Chetty, Friedman, and Saez 2013; Kleven and Waseem 2013; Kleven 2016; Kreiner, Munch, and Whitta-Jacobsen 2015; Labanca and Pozzoli 2022; Anagol, Davids, Lockwood, and Ramadorai 2022). While dynamic returns to effort represent a conceptually different mechanism, their existence may be driven by an underlying information friction: the fact that the verification of effort and productivity is costly to employers. As we show, such verification costs give rise to an equilibrium with intermittent performance evaluations and dynamic returns. This insight is related to career-concern models (Harris and Holmström 1982; Holmström 1999) in which employers have imperfect information about worker productivity, and implicit contracts link current effort to future wages. Our model captures similar ideas in a simple manner and informs empirical work on labor supply responses.

Finally, our paper is linked to a large body of empirical work studying wage determination and careers. This includes papers that compare the implications of standard labor supply models and contract models for changes in earnings and hours over time (Abowd and Card 1987, 1989), arguing that the standard model fits the data poorly. It also includes papers that document the importance of job-to-job mobility for wage growth (e.g., Topel and Ward 1992; Farber 1999; Abowd, Kramarz, and Margolis 1999; Pavan 2011; Card, Heining, and Kline 2013; Card, Cardoso, and Kline 2016; Card, Cardoso, Heining, and Kline 2018). Consistent with these literatures, we take a contract view on employment relationships and emphasize the critical role of job switches for earnings dynamics.

The paper proceeds as follows. Section 2 develops our theoretical model of dynamic compensation. Section 3 describes the data. Section 4 presents descriptive evidence on earnings and hours-worked patterns over the lifecycle, verifying the predictions of the model. Section 5 presents quasi-experimental evidence on earnings elasticities, using job switchers to uncover the long-run macro elasticity. Section 6 concludes and discusses policy implications.

They argue that the welfare-relevant earnings elasticity in a superstar market is larger than in a standard labor market due to a job switching mechanism. When superstar workers are induced to provide greater effort through lower taxes, they anticipate being reassigned to a better job and this amplifies the incentive. We share the focus on job switching, but our model is otherwise different and highlights the importance of job switching effects for all salaried workers, not just superstars.

2 A Theoretical Model of Dynamic Compensation

2.1 Setting

We consider a population of infinitely-lived workers with heterogeneous and time-varying productivities n_t. In each period, workers derive utility from consumption (which depends on *realized* earnings z_t) and disutility from effort (which depends on *latent* earnings y_t), where realized and latent earnings may be misaligned due to dynamic returns to effort. Flow utility is specified as

$$u_t = (1-\tau)\, z_t - n_t v\,(y_t/n_t)\,, \tag{1}$$

where τ is the marginal tax rate. The productivity parameter is specified as $n_t = g\,(t) + \mu$, where $g\,(t)$ is a common, deterministic lifecycle component and μ is an idiosyncratic, random shock.

The quasi-linear utility specification in (1) is common in the tax literature (e.g., Diamond 1998; Kleven, Kreiner, and Saez 2009; Saez 2010). However, the literature has focused on the standard labor supply model where $z_t = y_t$, i.e. where effort choices translate immediately and frictionlessly into realized earnings. In this special case, assuming that $v\,(x)$ takes the isoelastic form $\frac{\eta}{\eta+1}x^{\frac{\eta+1}{\eta}}$, worker optimization gives the familiar expression $z_t = y_t = (1-\tau)^{\eta}\, n_t$. Here η is the elasticity of earnings with respect to the marginal net-of-tax rate and productivity n_t represents potential earnings at a tax rate of zero.

We relax the assumption that effort maps immediately into earnings. In our model, realized earnings z_t change only at job events (such as switches between occupations or firms), which occur with probability λ in any given period. These job events align realized earnings with latent earnings (effort). Hence, we have

$$z_t = \begin{cases} y_t & \text{with probability } \lambda \\ z_{t-1} & \text{with probability } 1-\lambda. \end{cases} \tag{2}$$

The idea is that worker effort is unobservable without a costly performance evaluation, resulting in an employment contract where effort is rewarded discretely and intermittently at job events (performance evaluations). We start by assuming that the switching probability λ is exogenous, but we later develop a generalization where the switching probability is endogenously set by firms facing effort verification costs. In either case, the value of λ determines the degree to which the return to effort is dynamic. The special case of $\lambda = 1$ corresponds to the standard labor supply model in which the return to effort is immediate. Conversely, if λ is small, the return to effort materializes far into the future in expectation.[12]

In this model, effort y_t is a choice variable and earnings z_t is a state variable. At time t, workers know z_{t-1} and n_t, and maximize expected lifetime utility with respect to current and future efforts. Denoting the discount factor by δ, the optimization problem can be written as

$$\max_{\{y_s\}_t^{\infty}} \quad \sum_{s=t}^{\infty} \delta^{s-t} \mathbb{E}\left[u_s | z_{t-1}, n_t\right], \tag{3}$$

subject to equations (1)-(2). The solution can be characterized as follows:

Proposition 1 (Optimal Effort). *Assuming $v(x) = \frac{\eta}{\eta+1} x^{\frac{\eta+1}{\eta}}$, the optimal choice of latent earnings (effort) is given by*

$$y_t = \left(\frac{\lambda}{1 - (1-\lambda)\delta} \cdot (1-\tau) \right)^{\eta} n_t \qquad \forall t, \tag{4}$$

where η is the Hicksean elasticity of effort with respect to the marginal net-of-tax rate $1 - \tau$.

Proof. See Appendix B.1. ■

Optimal effort takes the standard form except for the adjustment factor $\frac{\lambda}{1-(1-\lambda)\delta}$. This factor captures the effect of dynamic returns to effort. When $\lambda = 1$, the level of effort is the same as in the standard model. Introducing dynamic returns ($\lambda < 1$) has two counteracting effects on effort. On the one hand, a lower λ implies that effort returns have longer delays and this disincentivizes effort (reflected in the numerator of the adjustment factor). On the other hand, a lower λ implies that effort returns, when they do materialize, are expected to last longer and this incentivizes effort (reflected in the denominator of the adjustment factor). In the special case of $\delta = 1$, these two effects offset exactly and the level of effort is the same as in the standard model. Importantly, this point pertains to the level of *latent* earnings, whereas we are ultimately interested in the response of *realized* earnings to taxes. Even when $\delta = 1$, the model has very different predictions than the standard model.

Using equation (2), we can write average earnings at time t as a function of the average levels

[12] It is worth pointing out that our model is conceptually related to a large macro literature studying rigid prices and wages. This literature has developed models with time-dependent price adjustment rules (Taylor 1980; Calvo 1983), state-dependent price adjustment rules (Caplin and Spulber 1987; Caplin and Leahy 1991; Caballero and Engel 1991), and a combination of the two elements (Nakamura and Steinsson 2010). The earnings specification in (2) is a form of Calvo contract in which there is a constant probability of earnings adjustment, independently of the time since the last adjustment. Importantly, our objective — understanding how earnings and effort respond to taxes — is fundamentally different from the macroeconomic focus on nominal price rigidity and the impact of monetary policy.

of effort from time 0 and the initial level of average earnings at time 0. We have

$$\bar{z}_t = \lambda \sum_{s=0}^{t} (1-\lambda)^s \, \bar{y}_{t-s} + \left(1 - \lambda \sum_{s=0}^{t} (1-\lambda)^s\right) \bar{z}_{-1}, \tag{5}$$

where average earnings $\bar{z}_t$ equals a weighted average of historical efforts $\bar{y}_0, ..., \bar{y}_t$ and initial average earnings $\bar{z}_{-1}$, with weights that depend on λ. This equation also describes each individual's expectation at time 0 of earnings at time t.

The model has the following predictions regarding the variation in effort and earnings:

Proposition 2 (Effort and Earnings Predictions). *The dynamic compensation model* ($\lambda < 1$) *has the following predictions that differ from the standard labor supply model* ($\lambda = 1$)*:*

1. *Average earnings $\bar{z}_t$ depend on past effort choices $\bar{y}_{s<t}$, conditional on current effort $\bar{y}_t$.*
2. *The contemporaneous correlation between earnings z_t and effort y_t equals the per-period switching probability λ.*
3. *For each worker, the lifecycle profile of earnings z_t is discrete around job switches.*
4. *For each worker, the lifecycle profile of effort y_t is smooth around job switches, as long as productivity and taxes are smooth.*

Proof. (1) This follows from equation (5) derived above. (2) See Appendix B.2. (3) This follows directly from the specification in (2). (4) This follows from equation (4) in Proposition 1. ■

2.2 Earnings Responses to Taxes

The model has important implications for earnings responses to taxes and welfare measurement. To see this, consider a permanent change in the tax rate from time 0, assuming that the economy is initially in a steady state with constant average earnings, $\bar{z}_t = \bar{z}$.[13] The welfare effect of such a tax change can be understood by considering its effects on tax revenue, $R_t = \tau \bar{z}_t$. A change in the tax rate has a mechanical effect on revenue, $dM_t = d\tau \cdot \bar{z}_t$, and a behavioral effect on revenue, $dB_t = \tau \cdot d\bar{z}_t$. Defining the elasticity of earnings at time t as $\varepsilon_t^z \equiv \frac{d\bar{z}_t/\bar{z}_t}{d(1-\tau)/(1-\tau)}$, the ratio of behavioral

[13]This assumption implies that we disregard any systematic lifecycle trends in earnings (i.e., $g(t)$ is constant). This simplifies the analysis, but is not important for the substance of the results. We consider the general case in appendix, as discussed below.

to mechanical effects can be written as

$$dD_t \equiv dB_t / dM_t = \frac{\tau}{1-\tau} \cdot \varepsilon_t^z. \tag{6}$$

This is a standard formula for the marginal deadweight loss of taxation (see e.g., Saez, Slemrod, and Giertz 2012; Kleven 2021), where the earnings elasticity ε_t^z is a sufficient statistic, conditional on τ. The issue is that, with dynamic compensation, ε_t^z increases over time and the measured welfare effect dD_t therefore depends on the time horizon of the elasticity estimation. Most quasi-experimental approaches allow only for the estimation of short-run elasticities and welfare effects, but policy design depends on long-run (steady state) welfare effects. In fact, assuming that the social planner puts equal weights on welfare now and in the future, the present value of social welfare is equivalent to steady state welfare.[14]

Computing the long-run welfare effect, dD_∞, requires information about the long-run earnings elasticity, ε_∞^z. To show how such information might be obtained empirically, we derive the following properties of earnings elasticities.

Proposition 3 (Earnings Elasticities). *Consider a permanent change in τ from time $t = 0$, assuming that the economy is initially in a steady state with constant average earnings, $\bar{z}_t = \bar{z}$. In this case, we have*

1. *The long-run elasticity of realized earnings with respect to the net-of-tax rate equals $\varepsilon_\infty^z = \eta$, i.e. the elasticity of latent earnings (effort) characterized in equation (4).*

2. *The elasticity of realized earnings at time t is a downward-biased estimate of the long-run elasticity. Specifically,*

$$\varepsilon_t^z = \alpha_t \eta \ \text{ where } \ \alpha_t = \lambda \sum_{s=0}^{t} (1-\lambda)^s \leq 1, \tag{7}$$

   such that the elasticity starts at the short-run level $\varepsilon_0^z = \lambda\eta$ and increases gradually towards its long-run level $\varepsilon_\infty^z = \eta$.

3. *For workers experiencing their first post-reform job switch at time t, the elasticity of realized earnings reveals the long-run elasticity, i.e. $\varepsilon_t^z|_{J_t=1} = \eta$ where $J_t = 1$ is an indicator for having the first post-reform job switch at time $t \geq 0$.*

Proof. (1) This follows from equation (7) as $\alpha_\infty = 1$ regardless of λ. (2) Using that the initial steady state must have $\bar{z}_t = \bar{y}_t$, equation (5) implies that $\varepsilon_t^z = \lambda \sum_{s=0}^{t} (1-\lambda)^s \cdot \varepsilon_t^y$ where $\varepsilon_t^y \equiv \frac{d\bar{y}_t/\bar{y}_t}{d(1-\tau)/(1-\tau)}$.

[14] See Appendix B.3 for a proof.

We have $\varepsilon_t^y = \eta$ from equation (4), which gives the relationship in (7). (3) It follows directly from equation (2) that $z_t|_{J_t=1} = y_t$ and, therefore, $\varepsilon_t^z|_{J_t=1} = \eta$. ■

This proposition shows that, in a world with dynamic compensation ($\lambda < 1$), estimates of short-run elasticities ε_t^z underestimate the welfare-relevant, long-run elasticity η. The bias is increasing in the degree of dynamic compensation (inversely related to λ). However, the last part of the proposition shows that the long-run elasticity can be uncovered from short-run responses by restricting the sample to job switchers because, for such individuals, realized and latent earnings momentarily coincide. The next section investigates identification based on switchers in greater depth.

Finally, while the preceding derivations disregard lifecycle trends in earnings, the results are generalized to allow for such lifecycle dynamics in Appendix B.4. There we show that the formula for α_t becomes more involved, but it remains the case that it increases over time from $\alpha_0 = \lambda$ to $\alpha_\infty = 1$.

2.3 Heterogeneity and Identification

The preceding analysis allowed for heterogeneity in earnings via the idiosyncratic productivity term μ, but all other parameters of the model were assumed to be homogeneous across workers. Realistically, there will also be heterogeneity in effort elasticities η and switching probabilities λ. Denoting the joint density of these two parameters by $f(\eta, \lambda)$, we are interested in estimating the average long-run earnings elasticity, i.e.

$$\mathbb{E}\left[\varepsilon_\infty^z\right] = \int_\lambda \int_\eta \eta f(\eta, \lambda)\, d\eta d\lambda = \mathbb{E}\left[\eta\right]. \tag{8}$$

Our ability to identify this macro elasticity using switchers will depend on the properties of $f(\eta, \lambda)$. We can estimate the average earnings elasticity among workers making their first post-reform job switch at time t, i.e.

$$\mathbb{E}\left[\varepsilon_t^z | J_t = 1\right] = \int_\lambda \int_\eta \eta f(\eta, \lambda | J_t = 1)\, d\eta d\lambda, \tag{9}$$

where $f(\eta, \lambda | J_t = 1)$ denotes the density of η, λ among such first-time switchers. The following proposition characterizes the conditions under which this estimand recovers the long-run parameter of interest.

Proposition 4 (**Identification**). *Consider a permanent change in τ from time $t = 0$. For workers making their first post-reform job switch at time $t \geq 0$, the average elasticity of realized earnings identifies*

$$\mathbb{E}\left[\varepsilon_t^z | J_t = 1\right] = \mathbb{E}\left[\eta\right] + \frac{\operatorname{cov}\left(\eta, \lambda\left(1-\lambda\right)^t\right)}{\mathbb{E}\left[\lambda\left(1-\lambda\right)^t\right]}, \tag{10}$$

where $\lambda\left(1-\lambda\right)^t \equiv P\left(J_t = 1|\lambda\right)$ is the probability of making the first post-reform job switch at time t for a worker of type λ. If $\eta \perp \lambda$, we have $\operatorname{cov}\left(\eta, \lambda\left(1-\lambda\right)^t\right) = 0$ and therefore $\mathbb{E}\left[\varepsilon_t^z | J_t = 1\right] = \mathbb{E}\left[\eta\right]$ for $\forall t \geq 0$.

Proof. From Bayes' Rule, we have $f\left(\eta, \lambda | J_t = 1\right) = \frac{P(J_t=1|\lambda) f(\eta,\lambda)}{P(J_t=1)}$, where $P\left(J_t = 1|\lambda\right) = \lambda\left(1-\lambda\right)^t$ and $P\left(J_t = 1\right) = \mathbb{E}\left[\lambda\left(1-\lambda\right)^t\right]$ denote conditional and unconditional probabilities of making the first post-reform job switch at time $t \geq 0$. Inserting this into equation (9), we obtain

$$\mathbb{E}\left[\varepsilon_t^z | J_t = 1\right] = \frac{\mathbb{E}\left[\eta \cdot \lambda\left(1-\lambda\right)^t\right]}{\mathbb{E}\left[\lambda\left(1-\lambda\right)^t\right]}. \tag{11}$$

Using the definition of covariance ($\operatorname{cov}\left(X, Y\right) = \mathbb{E}\left[XY\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$), this corresponds to the result in equation (10). ■

Hence, under orthogonality between η and λ, the long-run macro elasticity can be point identified. Importantly, this case is associated with an observable feature of the data, namely that the switcher elasticity $\mathbb{E}\left[\varepsilon_t^z | J_t = 1\right]$ is constant in t. We verify that this condition is satisfied in our empirical application, consistent with point identification. At the same time, because the condition may not hold across all settings, it is relevant to consider situations where η and λ are correlated. In such situations, partial identification is still possible. We have:

Corollary 1 (**Partial Identification**). *The probability of making the first post-reform job switch at time t, $P\left(J_t = 1|\lambda\right) = \lambda\left(1-\lambda\right)^t$, is increasing in λ for $t < \frac{1-\lambda}{\lambda}$ and decreasing in λ for $t > \frac{1-\lambda}{\lambda}$. Therefore, if $\operatorname{cov}\left(\eta, \lambda\right) > 0$, we have that $\operatorname{cov}\left(\eta, \lambda\left(1-\lambda\right)^t\right)$ is positive at $t = 0$ and turns negative at a sufficiently large t. From equation (10), this implies that short-run switcher elasticities, $\mathbb{E}\left[\varepsilon_t^z | J_t = 1\right]$ for small t, provide upward-biased estimates of the average long-run elasticity $\mathbb{E}\left[\eta\right]$. In this case, with estimates of short-run switcher elasticities for periods $t = 0, ..., T$, a lower bound is given by*

$$\sum_{t=0}^{T} \Lambda_t \mathbb{E}\left[\varepsilon_t^z | J_t = 1\right] = \sum_{t=0}^{T} \Lambda_t \mathbb{E}\left[\eta | J_t = 1\right] \leq \sum_{t=0}^{\infty} \Lambda_t \mathbb{E}\left[\eta | J_t = 1\right] = \mathbb{E}\left[\eta\right], \tag{12}$$

where Λ_t denotes the share of workers making their first post-reform job switch at time t. Conversely, if

$\text{cov}(\eta, \lambda) < 0$, *then* $\text{cov}\left(\eta, \lambda(1-\lambda)^t\right)$ *is negative at* $t = 0$ *and turns positive at a sufficiently large* t. *From equation (10), this implies that* $\mathbb{E}\left[\varepsilon_0^z | J_0 = 1\right] < \mathbb{E}\left[\eta\right]$ *is a lower bound.*

Proof. These results follow from Proposition 4 by noting that $\frac{dP(J_t=1|\lambda)}{d\lambda} = (1-\lambda)^t \left(1 - t\frac{\lambda}{1-\lambda}\right)$. ■

To conclude, in a world with dynamic compensation, short-run switchers can be used to either point identify or partially identify the long-run macro elasticity of interest.

2.4 Endogenous $\boldsymbol{\lambda}$

Appendix B.5 develops an extension of our model with an endogenous switching probability λ. In this model, worker effort is unobservable without a performance evaluation. The cost of evaluating a given worker is q and reveals true effort in the current period. Evaluations are carried out randomly with frequency λ. The equilibrium is given by the constrained-efficient solution in which chosen effort and evaluation frequency maximize worker-firm surplus. In a steady state, this amounts to maximizing

$$S = (1-\tau)\left[y - q\lambda\right] - nv\left(y/n\right), \tag{13}$$

where we assume that evaluation costs $q\lambda$ are tax deductible. This will be the case if, for example, the costs of performance evaluations reflect verifiable labor or equipment costs.

In this framework, we obtain the following proposition.

Proposition 5 (Endogenous $\boldsymbol{\lambda}$). *With a positive and finite evaluation cost* q*, the equilibrium evaluation frequency* $\lambda \in (0,1)$*. The limit case of perfect verification* ($\lambda = 1$) *is obtained for* $q = 0$*, while the limit case of no verification* ($\lambda = 0$) *is obtained for* $q = \infty$*. Outside these limit cases,* λ *is decreasing in the evaluation cost* q*, increasing in the effort elasticity* η*, decreasing in the discount factor* δ*, and independent of* τ*.*

Proof. See Appendix B.5. ■

This proposition microfounds the dynamic compensation model and implies that all of our previous results generalize. Two points are worth highlighting. First, the switching probability is independent of the tax rate due to the assumption that evaluation costs are tax deductible. With partial or no deductibility, there would be an impact of taxes on the probability of switching. We will directly test for this in our quasi-experimental analysis. Second, in the extended model, there will be heterogeneity in λ if there is heterogeneity in evaluation costs q and/or in effort elasticities η. The identification results of the previous section depend on the correlation between η and λ, which will be governed by the joint density of (η, q) and the characterization of λ as a function of q and η in Proposition 5.

2.5 Other Extensions

2.5.1 Competitive Labor Market Equilibrium

Given the wage dynamics in equation (2), if worker productivity n_t grows over time, output y_t is on average higher than wages z_t. This is inconsistent with a competitive equilibrium in which firms make zero expected profits. It is possible to generalize our results to a competitive equilibrium where average wages equal average output.

Suppose productivity n_t grows at a constant rate g. If a worker experiences a job switch at time t, the wage of the worker is adjusted to a competitive market level, denoted by $\hat{z}_t$, and stays fixed until the next job switch. This implies that equation (2) is replaced by

$$z_t = \begin{cases} \hat{z}_t & \text{with probability } \lambda \\ z_{t-1} & \text{with probability } 1-\lambda. \end{cases} \tag{14}$$

In Appendix B.6, we show that the equilibrium level of latent earnings (effort) is given by

$$y_t = \left(\frac{\lambda}{1-\delta\left(1-\lambda\right)\left(1+g\right)} \cdot \left(1-\tau\right) \right)^{\eta} n_t, \tag{15}$$

and that the realized earnings of job switchers equal

$$\hat{z}_t = y_t \frac{1-\delta\left(1-\lambda\right)}{1-\delta\left(1+g\right)\left(1-\lambda\right)}. \tag{16}$$

These equations imply that, without productivity growth ($g = 0$), allowing for labor market competition and zero expected profits have no impact on effort and earnings. With positive growth, the earnings level of job switchers is higher than their contemporaneous effort ($\hat{z}_t > y_t$) to reflect that their output is expected to increase over time. This also leads to a higher level of effort y_t. Importantly, these level effects on effort and earnings do not change any of the main results. In particular, the structural effort elasticity is still equal to η, and a given percentage change in effort still translates into the same percentage change in realized earnings among job switchers. These are the key properties needed for the remaining results to go through.

2.5.2 Introducing Earnings Dynamics for Job Stayers

We have assumed that effort and earnings are completely unrelated absent a job switch. However, this assumption is not needed for identification. Consider replacing equation (2) with

$$z_t = \begin{cases} y_t & \text{with probability } \lambda \\ \theta y_t + (1-\theta)\, z_{t-1} & \text{with probability } 1-\lambda, \end{cases} \tag{17}$$

where $\theta = 0$ corresponds to the baseline model. We show in Appendix B.7 that, in this case, the optimal choice of effort equals

$$y_t = \left(\frac{\lambda + (1-\lambda)\,\theta}{1-(1-\lambda)\,(1-\theta)\,\delta} \cdot (1-\tau) \right)^{\eta} n_t \qquad \forall t. \tag{18}$$

As before, the structural effort elasticity equals η and corresponds to the long-run elasticity of realized earnings. This may be compared to the short-run elasticity of realized earnings, derived from equations (17)-(18):

$$\varepsilon_t^z = \alpha_t \eta \ \text{ where } \ \alpha_t = (\lambda + (1-\lambda)\,\theta) \sum_{s=0}^{t} [(1-\lambda)\,(1-\theta)]^s \leq 1, \tag{19}$$

which provides a downward-biased estimate of η whenever $\theta < 1$. Most importantly, it follows from equations (17)-(18) that the short-run elasticity for job movers continues to identify the long-run elasticity, i.e. $\varepsilon_t^z|_{J_t=1} = \eta$.

2.5.3 Generalized Earnings Dynamics for Movers and Stayers

We have assumed that effort and earnings become perfectly aligned whenever there is a job switch. In reality, the evaluation process for job movers may be imperfect, for example due to a delayed link between effort and observable output. We allow for such aspects by replacing equation (2) with

$$z_t = \begin{cases} \theta^m y_t + (1-\theta^m)\, z_{t-1} & \text{with probability } \lambda \\ \theta^s y_t + (1-\theta^s)\, z_{t-1} & \text{with probability } 1-\lambda. \end{cases} \tag{20}$$

We allow for the evaluation of job movers to be imperfect ($\theta^m < 1$), while maintaining that it provides more information than what is available for job stayers ($\theta^m > \theta^s$). This specification is analyzed in Appendix B.8. The long-run earnings elasticity still equals η, but this parameter is no longer fully captured by the short-run elasticity for job movers due to $\theta^m < 1$. The short-run elasticity for movers, $\theta^m \eta$, provides a lower bound for the long-run elasticity. Importantly, it is less downward-biased than the standard short-run elasticity based on pooling movers and stayers, i.e. $\lambda\theta^m \eta + (1 - \lambda)\,\theta^s \eta$.

3 Data

The empirical analysis is based on administrative data covering the full population of Denmark from 1980 to 2018. The data combine several administrative registers (linked at the individual level via personal identification numbers) and contain detailed information on earnings, hours worked, occupation, firm, and demographic variables. Virtually all of the information in the data is third-party reported (see Kleven, Knudsen, Kreiner, Pedersen, and Saez 2011).

Two features of the data are worth highlighting. First, the data are employer-employee matched and include detailed occupation codes, allowing us to observe jobs (firm$\times$occupation cells) at a granular level.[15] The occupation codes build on the International Standard Classification of Occupations (ISCO), adapted by Statistics Denmark and called DISCO codes. The classification system has changed over time. Since 2010, occupations have been coded according to the DISCO-08 classification (563 occupations), while between 1991-2009, occupations were coded according to the DISCO-88 classification (372 occupations). We bridge this data break using a crosswalk developed by Humlum (2021). Prior to 1991, occupation codes were based on an older Danish classification system (299 occupations). As this classification is still available after 1991, we are able to bridge the old occupation codes with the more recent ones. As a rule, private employers with at least 10 workers and all public employers must register and report the occupation of each worker to Statistics Denmark. For the remaining workers, Statistics Denmark imputes occupation codes based on industry, labor union, and education. Table A.1 in the appendix shows examples of top and bottom occupation titles, ranked by average wage earnings.

Second, the data include two administrative measures of hours worked. We are ultimately interested in the role of dynamic returns to *effort*, which is conceptually different from time spent at work.[16] Still, given hours worked is a component of effort, we provide descriptive evidence

[15] Firm and occupation codes may sometimes change without an actual job switch (e.g., due to changes in ownership structure or reclassifications of worker groups). To avoid attributing such data changes to job switches, we drop all switches where more than 50% of an individual's coworkers move to the same job cell in the same year.

[16] In general, observed working hours may deviate from true effort for two reasons. One reason is that working hours reported in the data may reflect contracted hours rather than actual hours, or some mixture between the two. The other

on the relationship between hours and earnings that speaks to the predictions of the model. The first measure comes from a mandated pension scheme — *Arbejdsmarkedets Tillægspension* (ATP) — which requires employers to contribute on behalf of their employees based on individual working hours. The pension contribution is a function of a binned measure of working hours. Specifically, for someone paid monthly — the typical contract for salaried workers — the annual contribution depends on annual hours $\sum_{m=1}^{12} h_m$, where monthly hours h_m are divided into four bins.[17] This measure is available for the entire period 1980-2018, but has the disadvantage of being capped at full time for all 12 months of the year.[18] The second measure is better, but is only available since 2008. This measure provides information on uncapped hours for all workers at the monthly level. We use the first measure for analyses requiring a long panel and the second measure for analyses requiring us to capture hours variation precisely, including among full-time workers.

4 Descriptive Evidence on Dynamic Compensation

This section presents descriptive evidence that confirms the predictions of the theoretical model and motivates the quasi-experimental approach in the next section. While our evidence overlaps with existing findings in the labor literature, we leverage the uniquely rich Danish data to provide new insights on the role of past effort and job switches for earnings outcomes. We emphasize three empirical facts, all of which are consistent with our dynamic compensation model *and* inconsistent with standard labor supply models.[19]

4.1 Fact 1: Past Hours Worked Predict Current Earnings, Conditional on Current Hours

Any model with dynamic returns to effort predicts that earnings depend on past hours worked, even after conditioning on current hours worked. Figure 1 investigates if this prediction is borne out by the data. The figure is based on a balanced panel of workers observed between the ages of 20 and 50, showing how earnings at age 50 relate to current and past hours worked. Each panel shows the non-parametric relationship between the average earnings rank at age 50 and hours

reason is that unobserved effort choices influence the quality-adjusted hours relevant for earnings progression.

[17] This gives a total of 37 hours bins ($= 4 \times 12 - 12 + 1$) over a year.

[18] Even so, Appendix Figure A.1 shows that the administrative pension measure of hours worked aligns well with information from labor force survey data. The figure shows that the relationship between hours worked and earnings is very similar in the administrative and survey data. However, the survey data are much more noisy, especially at the top of the distribution.

[19] Specifically, the evidence presented below confirms the predictions listed in Proposition 2. These predictions cannot be explained by standard labor supply models, including dynamic human capital models.

worked at different ages. Panel A considers current working hours (at age 50), while Panel B considers past working hours (at ages 40-49). There is a strong positive relationship in both panels: working more hours, past or present, is associated with higher earnings. However, the fact that hours worked are correlated over time complicates the interpretation. The positive relationship between earnings and past hours may reflect that variation in past hours captures variation in current hours. Moreover, past hours may be correlated with skill parameters that impact earnings directly. The subsequent panels investigate if the predictive power of past hours is the result of such correlations.

Panel C shows the relationship between earnings rank at age 50 and hours worked between ages 40-49 without any controls (blue dots), with controls for current hours (orange dots), and with controls for current hours, demographics, occupation, and school grades (red dots).[20] The controls dampen the correlation between earnings and past hours as one would expect, but the relationship remains strong. Even with the full set of controls for current hours and skill variables, the expected earnings rank at age 50 increases from the 30th percentile to the 65th percentile as annual hours worked over the preceding 10 years increases from zero to 2,000 hours. The relationship is stronger at high levels of hours and earnings: making it to the top of the distribution requires consistently high effort over time. As a robustness check, Panel D considers the effect of past hours worked over a longer time horizon. This hardly affects the relationship.

This evidence suggests that the return to effort has a strong dynamic component. Our findings relate most directly to research on experience and tenure effects.[21] This literature documents correlations between wages and past employment or tenure, but it generally does not consider working hours and does not have such detailed proxies for skill confounders. In the online appendix, we supplement the analysis with evidence on the *contemporaneous* correlation between hours and earnings. This evidence shows that, while past hours are highly predictive of earnings, the correlation between current hours and earnings is very weak at the top of the distribution.[22] Our combined

[20] These controls absorb variation in skill or earnings capacity. The demographic controls include dummies for education level (8 categories), gender (binary), children (binary), and marital status (7 categories). Occupation dummies are based on 2-digit DISCO codes. School grades are measured as the GPA for the highest education degree obtained by age 21 (typically high-school GPA).

[21] See Willis (1986), Farber (1999), and Blundell and MaCurdy (1999) for reviews.

[22] Appendix Figure A.2 plots changes in log hours against changes in log earnings for workers in different parts of the earnings distribution: the bottom 20%, the top 20%, the top 10%, and the top 1%. The average relationship in each segment is depicted by blue dots, while examples of specific occupations are depicted by red triangles and diamonds. The figure shows a stark contrast between workers at the bottom and the top. At the bottom, the relationship between log hours and log earnings is very close to the 45-degree line, consistent with standard models of hourly-paid workers. At the top, the relationship between log hours and log earnings is virtually flat, consistent with our model of dynamic compensation.

evidence on the relationship between past hours, current hours, and earnings is consistent with our dynamic compensation model, but not with standard labor supply models.

4.2 Fact 2: Lifecycle Profiles in Earnings are Driven by Job Switches

We now turn to the role of job switching for the dynamics of earnings. The key idea of our paper is that the returns to effort are realized dynamically, at the point of job switches. Importantly, the objective in this section is not to investigate if job switches have causal effects on earnings. In our model, the causal driver of realized earnings is latent earnings, the profile of which reflects changes in effort and productivity over time. Job switches mediate the link between effort and earnings, but have no independent causal effect. Our objective is to verify if the prediction regarding the mediating role of job switches for earnings dynamics is supported by the data.

Leveraging the granularity of the Danish data, we measure job switches as transitions between firm×occupation cells. The first set of results is presented in Figure 2. Based on a balanced panel of workers observed between the ages of 20 and 50, the figure plots lifecycle profiles of earnings for different groups of workers. In Panel A, we compare workers in the top 10% and the bottom 50% of the earnings distribution at age 50. The two groups start at very similar earnings levels at age 20, but workers who make it to the top have a steeper lifecycle profile. The divergence in lifecycle profiles, illustrated in Panel B, leads to an earnings gap of about 1.7 log points at age 50. The question is how much of this divergence can be attributed to switches between job cells.

Panel C provides a striking answer. Starting from the raw difference in earnings profiles (dark blue), it shows the difference net of occupation fixed effects (light blue), net of occupation×firm fixed effects (orange), and net of occupation×firm×individual fixed effects (red). The impact of each set of controls depends on the order in which we include them, but we are ultimately interested in the total impact of including all of them. Theoretically, the discreteness of the lifecycle profile at job switches is a within-worker phenomenon, which is why we interact job fixed effects with individual fixed effects. The evidence shows that, within individual, job fixed effects explain almost all of the divergence in the lifecycle profiles of top and bottom earners. Within job cells, there is virtually no divergence between the two groups. In Panel D, we move further into the top tail of the distribution, comparing top 1% earners to bottom 50% earners. The results are very similar: more than 90% of the lifecycle divergence can be attributed to job transitions.

Do these results reflect that top earners make better switches or that they make more switches? Figure A.3 in the appendix shows that it is the former. The figure plots the distribution of the number of switches in the top and bottom samples analyzed above. The distributions are broadly similar in the different samples. The average number of switches is about 10 at the top and 9 at the bottom, corresponding to roughly one switch every three years. It is worth noting, however, that switching activity is not evenly spaced over the lifecycle. As workers age and reach higher earnings levels, switching becomes less frequent.

Another way of analyzing the importance of job switches is by decomposing the variance in earnings over the lifecycle into between-job variance and within-job variance. To implement this variance decomposition, note that the earnings of individual i in job j at time t can be written as

$$z_{ijt} = \bar{z}_{ij} + (z_{ijt} - \bar{z}_{ij}), \tag{21}$$

where $\bar{z}_{ij}$ denotes the average earnings of individual i in job j. Demeaning by the average earnings of individual i, $\bar{z}_i$, and taking variances gives

$$\text{var}\,(z_{ijt} - \bar{z}_i) = \underbrace{\text{var}\,(\bar{z}_{ij} - \bar{z}_i)}_{\text{Between Job}} + \underbrace{\text{var}\,(z_{ijt} - \bar{z}_{ij})}_{\text{Within Job}}, \tag{22}$$

where we use that $\text{cov}\,(\bar{z}_{ij} - \bar{z}_i, z_{ijt} - \bar{z}_{ij}) = 0$.

Figure 3 presents the results of such an analysis. Considering the same panel of workers as above, the figure plots the total variance of earnings (blue), the between-job variance (red), and the within-job variance (yellow) by earnings percentile at age 50. At all percentiles shown, virtually all of the lifecycle variation in earnings can be attributed to between-job variation. Consistent with the lifecycle graphs, between-job variance accounts for more than 90% of total variance. This holds at all earnings levels between the 50th and the 100th percentile. The robustness of the decomposition to the level of income suggests that this is a general feature of job contracts among salaried workers, who dominate a broad segment of the earnings distribution.

These results imply that job switches are central to understanding earnings dynamics and, by extension, to estimating earnings responses to taxes. Standard empirical approaches are not plausible in a world where all of the action is concentrated at switches that happen only intermittently.

4.3 Fact 3: Earnings Increase Discretely at Promotions, with No Change in Hours

The last piece of descriptive evidence focuses on earnings and hours changes around promotion events. In our model, positive job events — events where latent earnings are higher than current earnings — lead to sharp increases in earnings, with no change in effort. We verify this prediction based on an event study of within-firm promotions. Defining promotion as a switch to an occupation cell in which median earnings are at least 10% higher, we compare the outcomes of promoted and unpromoted co-workers over time. The results are robust to considering other promotion thresholds such as 5% or 20%.

To conduct the analysis, we use monthly data on earnings and hours worked, aggregated to the quarterly level. We match each promoted worker to their unpromoted co-workers within the same firm, giving unpromoted workers the same event time.[23] Letting Y_{iq} be the outcome of individual i in quarter q, indexed such that $q = 0$ is the first quarter of promotion, the event study regression is specified as

$$Y_{iq} = \boldsymbol{\alpha'} \boldsymbol{D_q^{Event}} + \beta D_i^{Treat} + \boldsymbol{\gamma'} \boldsymbol{D_q^{Event}} \times D_i^{Treat} + \phi_{q \in t} + \phi_a + \nu_{iq}, \tag{23}$$

where $\boldsymbol{D_q^{Event}}$ is a vector of quarterly event time dummies excluding the dummy for $q = -1$, D_i^{Treat} is a promotion dummy, $\phi_{q \in t}$ is a calendar year fixed effect, and ϕ_a is an age fixed effect. We include year and age fixed effects to neutralize time and lifecycle trends unrelated to promotions. The coefficients of interest are $\gamma_q \in \boldsymbol{\gamma}$. These are difference-in-differences coefficients that capture the effect of promotion in quarter q relative to the pre-promotion quarter $q = -1$ for promoted relative to unpromoted co-workers.

Figure 4 plots the difference-in-differences coefficients by event time for earnings (Panel A) and hours worked (Panel B). We see sharp effects on earnings: pre-trends are parallel, promoted workers experience a jump of about 4% at the time of the event, and the effect is stable over time. Conversely, there are no such effects on hours worked, which are smooth around the time of promotion. While these series have been normalized to zero at event time -1, it should be noted that there are level differences between promoted and unpromoted workers. Workers who get promoted tend to have higher working hours and earnings leading up to the event, consistent with the idea that promotions reward past effort.

Harking back to comments made in the previous section, these event studies should not be interpreted as estimating a causal effect of promotions on earnings. Viewed through the lens of our model, realized earnings are ultimately driven by effort and productivity, mediated through job switches due to the structure of job contracts. Saying that promotions are the reason for earnings

[23] In selecting the sample, individuals are required to stay in the same firm from two quarters before promotion to two quarters after promotion.

jumps corresponds to saying, for example, that tenure decisions are the reason for changes in academic salaries. This is true only in a narrow sense. The real reason is the quality of the academic CV, the returns to which are materialized at discrete tenure events. Our model and evidence imply that this is a general feature of job contracts, where workers cannot influence earnings through effort without a discrete job event.

5 Estimating Earnings Elasticities with Dynamic Compensation

5.1 A Quasi-Experimental Approach Using Job Switchers

We now turn to the study of long-run earnings responses to taxes. To obtain exogenous variation in tax rates, we use a major tax reform implemented in Denmark in 2009-10.[24] Prior to the reform, income was taxed according to a progressive schedule with three brackets, commonly referred to as the bottom, middle, and top brackets. The 2009 reform eliminated the middle bracket and raised the top bracket threshold. The implication of these policy changes was that taxpayers above an income threshold experienced lower marginal tax rates, while those below the threshold were largely unaffected. The threshold that separates treatments and controls was located at around the 70th percentile of the income distribution.

The 2009 tax reform is illustrated in Figure 5. Panel A shows the evolution of marginal tax rates for taxpayers above and below the treatment threshold. The reform-induced reduction in the top marginal tax rate was large, about 11 percentage points on average.[25] There was heterogeneity in the tax rate reduction depending on location within the treated income interval. This is shown in Panel B, which plots changes in the marginal net-of-tax rate $1 - \tau$ by taxable income bin. The net-of-tax rate increased by about 17 percent over most of the treated income interval, but the increase was even larger in a range just above the treatment threshold.

While the tax cut was relatively large, there is otherwise nothing unique about this experiment: it creates tax variation by income level of the sort typically used in the literature on behavioral responses to taxes (see Saez, Slemrod, and Giertz 2012). Indeed, the objective is to demonstrate our approach using a widely available source of tax variation. However, a concern with using the 2009 reform is that it coincides with the Great Recession, posing identification threats from heterogeneous effects of the recession and recovery. To address such concerns, we also consider a

[24] We refer to Jakobsen and Søgaard (2022) for a detailed description of the Danish tax system and the 2009 reform.

[25] To be exact, the average reduction in the marginal tax rate was 11.43pp above the threshold and 1.70pp below the threshold. The difference in tax treatment was therefore 9.73pp.

historical tax reform — the 1987-reform — which reduced marginal tax rates at the top relative to the bottom.

Reduced-Form Approach: Typical quasi-experimental studies of earnings responses to tax reform compare treated and untreated workers in a short time window around a reform. As shown by our theoretical model, such approaches underestimate long-run responses in a world where the returns to effort are dynamic and mediated by job switches. In general, the feasible time window in quasi-experimental studies is too short to allow most workers to switch jobs and realize the return to higher effort. The attenuation bias depends on the degree of job mobility; it will be stronger in populations where job switching is less frequent. Our model shows that the long-run earnings response can be point identified from short-run switchers if the timing of switching is not selected on the underlying structural elasticity, and that it can be partially identified in the presence of selection. The model also shows that the time profile of the observed earnings elasticity for switchers (within a short window) can be used to separate the cases with and without selection.

Our empirical strategy starts from a difference-in-differences specification of the form

$$\Delta \log z_i = \boldsymbol{\beta} \boldsymbol{D_i^{Treat}} \times D_i^{Mover} + \boldsymbol{\gamma} \boldsymbol{D_i^{Treat}} + \mu D_i^{Mover} + \nu_i, \tag{24}$$

where $\Delta \log z_i$ denotes the log change in the labor income of individual i.[26] As shown in Figure 5, individuals were treated differently by the tax reform depending on their baseline taxable income prior to the reform.[27] Hence, we divide baseline taxable income into a set of discrete bins, $\boldsymbol{D_i^{Treat}}$, using boldface to denote vectors and omitting a bin below the treatment threshold. We interact these baseline income bins with an indicator for switching jobs, D_i^{Mover}, equal to one if the individual moves between firm$\times$occupation cells after the reform. This allows us to estimate difference-in-differences coefficients separately for job stayers ($\gamma_b \in \boldsymbol{\gamma}$) and job movers ($\beta_b + \gamma_b \in \boldsymbol{\beta} + \boldsymbol{\gamma}$) in each income bin b.

Given the specification assigns treatment status based on pre-reform income, the main threat to identification is the presence of non-tax effects on earnings growth $\Delta \log z_{it}$ that vary by pre-reform income level. The most obvious confounder is mean reversion, as discussed extensively in the

[26] The outcome variable z_i equals total labor compensation, including contributions to employer-administrated retirement savings plans. This ensures that our estimates are not influenced by shifting between taxable compensation and tax-deferred retirement savings contributions.

[27] The measure of taxable income that determines treatment status includes labor income, transfers, pensions, alimony, and certain capital income items.

literature (see Saez, Slemrod, and Giertz 2012). If income consists of both permanent and transitory components, those with high pre-reform incomes tend to be selected on positive transitory shocks, creating downward bias in the estimated responses to lower taxes as the transitory shocks dissipate over time. Following Jakobsen and Søgaard (2022), we address this issue by running our regression separately in pre-reform and post-reform datasets. The pre-reform specification considers earnings growth between 2006-08 by 2006 income bin. The resulting placebo estimates capture the effects of non-tax confounders assuming these are stable over time, which we verify in the data. The post-reform specification considers earnings growth between 2008-10 (and later) by 2008 income bin. Denoting the placebo estimates by superscript P, we may estimate behavioral responses based on a triple-differences approach: $\gamma_b - \gamma_b^P$ for stayers and $\beta_b + \gamma_b - \left(\beta_b^P + \gamma_b^P\right)$ for movers. We may alternatively consider a quadruple-differences approach by comparing the triple-differences estimates for movers and stayers, i.e. $\beta_b - \beta_b^P$. In fact, our preferred estimates will be based on the quadruple-differences approach. As we shall see, netting out the estimate for stayers provides are more robust way of controlling for mean reversion, especially when estimating behavioral responses over longer time windows.[28]

Earnings Elasticities: We convert the reduced-form estimates of earnings responses to the 2009 tax reform into elasticities with respect to $1-\tau$. This is done using the quadruple-differences approach just described. Merging the pre-reform (placebo) and post-reform datasets, the earnings elasticity ε is estimated from the following regression:

$$\begin{aligned} \Delta \log z_{it} \quad = \quad & \varepsilon \cdot \Delta \log\left(1-\tau_{it}\right) \times D_{it}^{Mover} + \boldsymbol{\gamma} \boldsymbol{D_{it}^{Treat}} + \mu D_{it}^{Mover} \\ & + \left(\boldsymbol{\beta^P} \boldsymbol{D_{it}^{Treat}} \times D_{it}^{Mover} + \boldsymbol{\gamma^P} \boldsymbol{D_{it}^{Treat}} + \mu^P D_{it}^{Mover}\right) \times D_t^P + \nu_{it}, \end{aligned} \tag{25}$$

where D_t^P is a dummy equal to one for the placebo period. To understand this specification, note that the first line corresponds to equation (24) where the treatment indicator in the interaction term has been replaced by the log change in the net-of-tax rate $\Delta \log\left(1-\tau_{it}\right)$. By itself, this line would be a triple-differences estimate comparing the difference-in-differences for movers and stayers. The

[28]It is worth spelling out the identification assumptions of the triple-differences and quadruple-differences specifications, respectively. The triple-differences specification is based on the assumption that any non-tax difference in earnings growth above and below the treatment threshold (for example due to mean reversion) is constant over time. This allows us to correct for non-tax confounders by doing the difference-in-differences in pre-reform data. The quadruple-differences specification is based on the assumption that job stayers cannot respond to taxes and therefore provide additional information on non-tax confounders. Assuming no stayer responses, the specification allows for non-tax differences in earnings growth above and below the treatment threshold to be changing over time, as long as they are changing in the same way for movers and stayers.

second line corresponds exactly to equation (24) but is turned on only for the pre-reform (placebo) period. The combination of the two lines gives our quadruple-differences estimate of the earnings elasticity for job switchers.

A well-known issue with elasticity specifications like (25) is that, in a progressive tax system, the marginal tax rate τ_{it} depends on the choice of earnings z_{it}, creating bias from reverse causality. To address this issue, we follow the standard Gruber-Saez approach (Gruber and Saez 2002) of creating an instrument for $\Delta \log \left(1 - \tau_{it}\right)$ based on simulated mechanical tax changes $\Delta \log \left(1 - \tau_{it}^{sim}\right)$. The simulated tax changes account for changes in the tax code, but hold individual income choices fixed at their baseline values. Equation (25) is estimated based on 2SLS using $\Delta \log \left(1 - \tau_{it}^{sim}\right)$ as an instrument.

Effects on Job Switching Probability: While our main focus is on estimating earnings responses conditional on job switching, tax reforms may also affect the probability of switching itself. We studied such effects theoretically in section 2.4. Effects of taxation on the probability of switching jobs are interesting in their own right and, as discussed below, important for evaluating the possibility of selection bias. We estimate switching effects based on a difference-in-differences specification:

$$D_{it}^{Mover} = \boldsymbol{\beta D_t^{Year}} \times D_{it}^{Treat} + \boldsymbol{\mu D_t^{Year}} + \gamma D_{it}^{Treat} + \nu_{it}, \tag{26}$$

where D_{it}^{Mover} is a dummy equal to one if individual i made a job switch between year t and a base year $t - \Delta t$, $\boldsymbol{D_t^{Year}}$ is a vector of year dummies, and D_{it}^{Treat} is a treatment dummy equal to one if the individual's baseline income is above the treatment threshold. Equation (26) is a standard event study specification in which the parallel-trend assumption can be assessed based on the estimated coefficients $\hat{\beta}_t$ in the pre-reform years.

Selection into Switching: For the switcher elasticity to point identify the long-run elasticity in the population, the timing of switching cannot be selected on the underlying structural elasticity. To investigate such selection, we start by estimating if the switcher elasticity varies by the timing of switching jobs following the tax reform. Are earnings elasticities different for those who switch immediately after the reform compared to those who switch later? We find that the elasticity is very stable over time. This is consistent with orthogonality between switching probabilities and structural elasticities in which case the long-run elasticity can be point identified.

Still, a skeptic may be concerned that the stability of the switcher elasticity is the result of off-setting effects from selection and other time-varying confounders. We provide several additional analyses to address such concerns. The most natural way for selection to arise is that some workers switch jobs *because of the reform* and that such tax-induced switchers are selected on the elasticity. We can use our estimate of the effect of the tax reform on the probability of switching to bound this form of selection bias.

We denote the share of tax-induced switchers by $s_t = \Delta P_t / P_t$, where ΔP_t is the effect of the reform on the switching probability (estimated from equation 26) and P_t is the observed probability of any post-reform switch by year t, tax-induced or otherwise. Our estimate of the earnings elasticity among job switchers can be expressed as a weighted average:

$$\mathbb{E}\left[\varepsilon_t^z | J_t = 1\right] = s_t \cdot \eta^T + (1 - s_t) \cdot \eta^N, \tag{27}$$

where η^T and η^N are the structural elasticities for tax-induced switchers (T) and non-tax switchers (N). If $s_t > 0$ and $\eta^T \neq \eta^N$, our estimates will be biased by selection. The issue is not that tax-induced switchers are part of the estimate, but that they are over-represented: they are weighted by s_t — their share among short-run switchers — instead of their share in the full population ΔP_t. To bound the plausible amount of selection bias, we may consider a wide range of assumptions about the elasticity for tax-induced switchers, $\eta^T = \tilde{\eta}^T$. From equation (27), we obtain

$$\tilde{\eta}^N = \frac{\mathbb{E}\left[\varepsilon_t^z | J_t = 1\right] - s_t \cdot \tilde{\eta}^T}{1 - s_t}. \tag{28}$$

Now, for each elasticity pair $\left(\tilde{\eta}^T, \tilde{\eta}^N\right)$, we can calculate the implied population-wide elasticity based on the correct weights ΔP_t and $1 - \Delta P_t$. By allowing for extreme assumptions about $\tilde{\eta}^T$, we evaluate if selection into switching — coming from the effect of the tax reform on the probability to switch — could conceivably be an issue. We find that this is not a quantitatively important threat to identification. The reason is that the impact of taxes on the switching probability is too small to drive non-trivial selection bias.

We provide additional tests to address selection that is not driven by reform-induced switching. As shown by our theoretical model, selection may be an issue even when switching probabilities do not respond to the tax system. One of the tests we provide is to restrict the sample to plausibly exogenous switches, namely those triggered by mass layoffs. Conducting our quasi-experimental study of earnings responses in the sample of mass-layoff switchers, we find similar responses as in the full sample of switchers.

External Validity: Finally, we examine if our results are affected by the particular macroeconomic environment during the Great Recession by considering historical tax reforms.[29] There were four major income tax reforms in Denmark between the beginning of our data and the 2009 tax reform. Appendix Figure A.4 shows log-changes in the marginal net-of-tax rate by income level for each of these reform. Given we are interested in the elasticity at the top of the income distribution, among salaried career workers, we need a reform that changes the tax rate at the top relative to the bottom. As shown by the figure, only the 1987-reform fits the bill. This reform was also very large which is another empirical advantage. Hence, we investigate the external validity of our estimates by applying the switcher approach to the 1987-reform. As we shall see, the estimates are very similar for the two reforms.

5.2 Impact of Tax Reform: Switchers vs Non-Switchers

5.2.1 Reduced-Form Effects

We start by investigating the short-term earnings responses to the 2009 tax reform. Figure 6 shows results for all workers (Panel A) and for job movers vs job stayers (Panel B). Job movers are defined as workers who switch between firm×occupation cells. The figure plots changes in log earnings between 2008-10 (reform period) and between 2006-08 (pre-reform, placebo period) by baseline income bin, depicting the threshold that separates treatments and controls by a vertical line. As discussed above, the placebo series capture non-tax effects on earnings growth that vary by income level. This includes mean reversion: those with high baseline incomes tend to be selected on positive transitory shocks, reducing their earnings growth over time regardless of the tax cuts. Comparing the 2008-10 and 2006-08 series (above vs below the treatment threshold) gives a triple-differences estimate that controls for mean reversion.

Consider first the results for all workers in Panel A. The 2008-10 series shows that earnings growth is declining in baseline income level. As a result, a difference-in-differences approach that compares earnings growth above and below the treatment threshold would yield a negative effect. As can be seen from the 2006-08 series, this reflects the aforementioned bias from mean reversion. This series shows that, before the reform, there was a similar negative relationship between earnings growth and baseline income level. The two series track each other almost perfectly below the

[29] The Great Recession could bias our estimates if the effects of the recession and recovery vary between job movers and job stayers in a way that is correlated with the identifying tax variation. We show that the empirical patterns around the 2009 tax reform appear inconsistent with such macro confounders — including the fact that the estimates are stable during the recession *and* recovery — but it is difficult to rule out macro effects conclusively.

treatment threshold — consistent with mean reversion being stable over time — but diverge above the threshold. This divergence represents the causal effect of the tax reform on earnings. The effect is visually clear and precisely estimated, but modest in size. Reducing the top marginal tax rate by about 10pp increases earnings by less than 2%.

Consider now the results for job movers vs job stayers in Panel B. The modest response in the full sample masks strong heterogeneity by job switching status. For job movers, the pre-reform and post-reform series track each other perfectly below the treatment threshold and diverge sharply above the threshold. The earnings responses are large — about three times larger than in the full sample — and increasing in baseline income. For job stayers, on the other hand, we see no positive earnings responses to lower taxes. In fact, the post-reform series lies marginally below the pre-reform series above the treatment threshold, corresponding to a small negative effect. This seems to reflect an imperfection in the mean reversion adjustment for this subsample: the pre-reform series is rotated relative to the post-reform series, creating a small negative divergence in the treated range. The fact that the rotation occurs across both the treated and untreated regions suggests that the true response among job stayers is zero rather than negative.

The preceding analysis pools job movers regardless of whether they switch between firms or occupations. To see if the type of switch matters, Figure A.5 in the online appendix reproduces the analysis for firm switchers and occupation switchers separately. The empirical patterns are similar in these subsamples. These results suggest that the mediating role of job switches is independent of the particular type of switch. This is consistent with our theoretical model in which any switching event is associated with a performance evaluation that realigns latent and realized earnings.

The combination of large responses among job movers and no responses among job stayers suggests that the standard approach provides strongly downward-biased estimates of the true long-run effect. The reason is that job stayers do not stay in their current jobs permanently; they will eventually become job movers themselves and reveal their tax-induced earnings response at that time.[30] The idea of our approach — using short-term movers to estimate long-term responses in the population — is that job stayers increase *latent* earnings by as much as job movers, but that their return is not realized until the time of moving. As shown in section 2, point identification of the long-run elasticity requires orthogonality between structural elasticities and switching probabilities. This is associated with an observable pattern in the data, namely that the earnings

[30] As shown in Appendix Figure A.3, all workers switch jobs — and typically many times — over their careers.

response by job movers is constant in the timing of switching.[31] To investigate this point, we now consider the time profile of earnings responses among job movers.

To estimate such a time profile, we note that the reduced-form evidence provided above represents intent-to-treat (ITT) effects. Comparing individuals in different pre-reform tax brackets, regardless of their actual post-reform bracket location, creates attenuation bias as people move across brackets after the reform. Such attenuation tends to increase over time, confounding the interpretation of the time profile in ITT effects. Therefore, to estimate a time profile that relates to our theoretical predictions, we need to estimate treatment-on-the-treated (TOT) effects. This is done by expressing equation (24) in terms of an indicator for actual post-reform tax bracket, using pre-reform tax bracket as an instrument. The equation is estimated based on 2SLS.

Figure 7 shows the time profile of ITT effects (Panel A) and TOT effects (Panel B) on earnings. The red series provide estimates of long-run responses using our dynamic approach. Each dot represents a quadruple-differences estimate based on comparing earnings growth above vs below the treatment threshold for job movers vs job stayers in a post-period relative to a pre-period.[32] The pre-reform part of the series (2002-2008) gives placebo estimates by comparing, for each year t, earnings growth between year $t-2$ and t to earnings growth between year $t-4$ and $t-2$. The post-reform part of the series (2009-2014) gives cumulative estimates by comparing, for each year t, earnings growth between 2008 and year t to earnings growth between 2006 and 2008. As can be seen from the figure, the pre-reform estimates are very close to zero — validating the empirical design — while the post-reform estimates are stable from year 2010 onwards. The estimates are smaller in 2009, which is natural given the reform was not yet fully phased in at that time. Leaving aside this partially treated year, the earnings responses among job movers do not depend on the timing of the switch. This pattern is consistent with point identification of the long-run earnings elasticity.

Can we identify the long-run elasticity using a standard approach by considering a sufficiently long post-reform period? To answer this question, Figure 7 shows time profiles of earnings responses based on two "standard approaches." Both of these capture average responses in the full sample of workers, including job stayers. The first approach (solid black line) is based on scaling

[31]Conversely, if structural elasticities and switching probabilities are correlated, the earnings response by switchers is either declining (positive correlation) or increasing (negative correlation) as we consider switches farther removed from the time of the reform.

[32]The specification corresponds to equation (25) where $\Delta \log(1-\tau_{it})$ is replaced with a treatment dummy $\boldsymbol{D_{it}^{Treat}}$. In the figure, we control for baseline job cell by including occupation×firm fixed effects, but we show later that the results are robust to adding or dropping controls.

the estimates for job movers (red line) using the cumulated fraction of workers who have made a post-reform job switch, as characterized by equation (7). This corresponds to estimating the mover response from our quadruple-differences specification while imputing the stayer response to zero. We see that the average response increases over time as more workers make job transitions, but the convergence between the standard and dynamic approaches is slow. The second approach (dashed black line) estimates earnings responses in the full sample using a Gruber-Saez approach, avoiding the imputation of stayer responses to zero. When running the estimation in the full sample, pooling movers and stayers, we have to use the triple-differences specification. As discussed above, this specification requires that mean reversion effects are constant over time. We see that the Gruber-Saez approach aligns well with the scaled-dynamic approach over the first three years, but then the two approaches begin to diverge. The Gruber-Saez series fall back towards zero, likely due to changes in mean reversion that the triple-differences approach is unable to handle.[33] Taken together, the results in Figure 7 demonstrate how difficult it is to estimate long-run responses based on standard approaches.

Finally, we ask if taxes impact the probability of switching jobs. Figure 8 provides event study evidence based on equation (26). The figure shows effects on all switches between occupations or firms (blue series), occupation switches (red series), and firm switches (yellow series). Before the reform, there is parallel trends in the switching probabilities of treatment and control groups. After the reform, there is a modest divergence in the switching probabilities of the two groups, coming from occupation switches. The effect of the reform on the probability of job switching is about 0.02 and precisely estimated.[34] We investigate below if this amount of reform-induced switching could create non-trivial selection bias.

5.2.2 Elasticities

In Table 1, we convert our estimates of earnings responses to the 2009 tax reform into elasticities with respect to the marginal net-of-tax rate. Panel A shows elasticities for job switchers based on specification (25), estimated by using simulated tax changes $\Delta \log \left(1 - \tau^{sim}\right)$ as an instrument for actual tax changes $\Delta \log \left(1 - \tau\right)$. The first stage is extremely strong, with an F-statistic of about 16,000 after two years and 14,000 after four years. The structural earnings elasticity obtained from

[33]This is a key reason why tax reform studies based on the Gruber-Saez approach are unable to estimate effects over long time windows. Gruber and Saez (2002) themselves estimated effects over at most three years.

[34]Recall that observing small (or zero) effects of taxes on the probability of job switching is consistent with our theoretical model. As shown in section 2.4, the magnitude of the effect depends on the tax deductibility of the costs of verifying true effort (latent earnings).

2SLS estimation equals 0.486 after two years and 0.487 after four years. Hence, the elasticity is extremely stable in the timing of job switching. Panel B compares the switcher elasticities to the elasticities obtained from a standard approach pooling all workers. As above, the standard approach is implemented by scaling the switcher elasticity using the cumulated fraction of workers who have made a post-reform job switch (after two and four years, respectively), thus assigning a zero elasticity to job stayers. Even after four years, the standard elasticity is only about half the size of the switcher elasticity.

These results suggest that conditioning on job switching yields earnings elasticities that are large *and* stable over time, consistent with point identification of the long-run elasticity. Table 2 investigates the robustness of this finding to alternative samples and specifications. The left columns show estimates of two- and four-year earnings elasticities, while the right columns show estimates of two- and four-year effects on the probability of job switching. Consider first the elasticity estimates. Panel A maintains our baseline specification, but explores heterogeneity by income level. Job switchers in the top decile of the distribution have elasticities that are even larger and, importantly, still stable over time. Panel B considers different control variables, allows for "donut hole" specifications, and restricts attention to different types of job switches.[35] The general takeaway is that our findings are robust: the earnings elasticity for job switchers is around 0.4-0.6 across the many specifications — similar to the baseline estimate of about 0.5 — and it is stable across the two- and four-year horizons.

Turning to the probability of job switching, the table confirms that the reduction of the top tax rate increased the amount of job switching in the years following the reform. The four-year effect on the switching probability equals 0.01-0.03 across most specifications, with a baseline probability of about 0.5. The effect is therefore modest, but robust and precisely estimated. These results are interesting in their own right, and they are relevant for evaluating selection bias. In the next section, we use the estimated effects on the switching probability to bound selection bias.

[35] The baseline specification includes fixed effects for baseline job cell (firm×occupation) and is estimated using all workers with baseline income between 250,000 DKK and 1,000,000 DKK. In the table, we show the implications of dropping all controls and of adding granular demographic controls, either on their own or together with job fixed effects. We also show the implications of donut-hole specifications in which we drop observations near the treatment threshold (a small donut of +/- 2.5% of the threshold and a large donut of +/- 5.0% of the threshold). Finally, while the baseline specification considers any type of job switch, the table provides estimates specifically for occupation switches, firm switches, and mass-layoff firm switches. We discuss the results for mass-layoff switches in the analysis of selection below.

5.3 Identification: Is Switching Selected?

The fact that the earnings responses for job switchers are stable over time (Figure 7) is consistent with orthogonality between structural elasticities and switching probabilities. This is the central feature of the data we need for point identification of the long-run macro elasticity. In this section, we present three additional analyses to alleviate concerns about selection bias.

Bounds on Selection Bias: The most natural source of selection bias is the presence of *reform-induced* switching. Workers who choose to switch jobs because of the reform may be selected on large elasticities. Because we have estimated the effect of the reform on the switching probability — finding a positive, but modest effect — we may quantify this selection effect under different assumptions about the underlying structural elasticity for tax-induced switchers. By allowing for extreme elasticities, we can bound the amount of possible selection bias. The details of the approach were described in section 5.1.

Figure 9 compares our baseline estimate of the earnings elasticity for job movers (black line) to the true earnings elasticity (red line) as a function of the elasticity among tax-induced movers.[36] We allow tax-induced movers to have elasticities of up to four, well beyond the conceivable range. As we can see, the amount of selection bias remains modest even under extreme elasticities. The simple reason is that the effect of the reform on the switching probability is too small to create non-trivial selection bias.

Impact on Switcher Characteristics: Even though the reform had a small effect on the amount of job switching, it could have changed the *composition* of switchers. The preceding analysis does not address such selection effects. To investigate this issue, we estimate the impact of the reform on the characteristics of job switchers, leveraging a quasi-experimental design similar to the one used for our main outcome variables. The results are presented in Appendix Figure A.6. It provides event studies of different variables, comparing switchers above and below the treatment threshold. We focus on six variables that are important for labor supply and earnings: age, gender, marital status, number of children, occupational rank, and firm size. For all six variables, the treatment and control series are virtually parallel over time, before *and* after the reform, consistent with no effect on the composition of switchers. The difference-in-differences estimates (displayed in the

[36] The figure is based on our *four-year* estimates of the earnings elasticity and effect on job-switching probability. Alternative time horizons give similar results.

figure) are very small, albeit statistically significant due to the statistical power of our data.[37]

Mass-Layoff Switchers: As a final identification check, we consider job switches triggered by a plausibly exogenous event: a mass layoff. A large literature on the effects of job displacement has used mass layoffs as a source of exogenous variation (e.g., Jacobson, LaLonde, and Sullivan 1993). Building on this literature, we implement our quasi-experimental approach in the sample of workers who switch firms following a mass layoff. The findings are presented in Figure 10. The figure plots earnings responses by income bin for the full sample of movers (Panel A) and for the sample of mass-layoff movers (Panel B). Mass-layoff movers are defined as workers who switch to a new firm, coming from a firm that reduced their workforce by at least 30% in the year of the switch.[38] We find that the earnings responses to lower taxes are large even among mass-layoff movers. In fact, the responses are *larger* than in the full sample of movers. Hence, the large switcher elasticities documented above do not appear to be driven by selection into switching.[39]

Taken together, the analyses presented in this section suggest that our estimates of switcher elasticities are not biased, or at least not *upward* biased, by selection into job switching.

5.4 External Validity: 1987 Tax Reform

Our approach to studying dynamic compensation effects among career workers requires exogenous variation in the taxation of top earners. The 2009 tax reform is useful because it created large variation in the top tax rate and is relatively recent. At the same time, this reform raises concerns about the macroeconomic environment around the time of the Great Recession. It is therefore important to investigate if our findings are robust to using other tax reforms. In the modern era, there has been four major tax reforms in Denmark prior to the 2009 reform. Among these, only the 1987 reform significantly changed the taxation of top earners relative to bottom and middle earners.[40]

[37] These findings suggest that demographic controls should not make any material difference to our quasi-experimental estimates of earnings responses. This is consistent with the results shown in Table 2 and in Appendix Figure A.7.

[38] This corresponds to the definition of mass layoffs in Jacobson, LaLonde, and Sullivan (1993). The definition is meaningful only for larger firms, so we further restrict the sample to firms with at least 20 employees at the time of the mass layoff. About one-fifth of job switchers in our baseline sample satisfy these mass-layoff criteria. Alternative definitions of mass layoffs give qualitatively similar results, but stricter definitions (increasing the minimum fraction laid off and/or the minimum number of employees) reduce sample size and increase standard errors.

[39] The positive earnings effect of lower taxes among mass-layoff movers is consistent with a negative earnings effect of the mass layoff itself, as documented in the literature on job displacement. To confirm this, Appendix Figure A.8 provides an event study of mass layoffs, showing that they generate sizeable and persistent earnings losses. Hence, the findings in Figure 10 should be interpreted as showing that, within the group of switchers affected by mass layoffs, those who received tax cuts were less negatively affected than those who did not receive tax cuts.

[40] See Appendix Figure A.4.

Several previous papers have analyzed this reform, estimating earnings elasticities of around 0.1 (see e.g., Jakobsen and Søgaard 2022). In Table 3, we provide estimates for the 1987 reform based on our dynamic switcher approach. The table is constructed exactly as Table 2 for the 2009 reform.

The findings for the 1987 reform are consistent with those for the 2009 reform. Considering our baseline specification and the full sample of job switchers, the earnings elasticity equals 0.477 after two years and 0.421 after four years, very close to our previous estimates. Looking across the large number of robustness checks, the results for the two reforms are qualitatively similar. The earnings elasticity is around 0.4-0.6 across most specifications. The main difference is that, for the 1987 reform, the elasticity is somewhat less stable over time in some of the specifications. The effect of the 1987 reform on the probability of job switching is very small, again consistent with our previous estimates.

To conclude, our results are externally valid to a historic Danish tax reform. The larger question is whether our results are also externally valid to other countries, including countries with very different labor market institutions and social norms. Because our approach relies on widely available tax reform variation, it would be relatively straightforward to implement in other countries. This will be an important agenda for future research.

6 Conclusion

The idea that the return to effort is dynamic seems *prima facie* true, especially for career workers at the top of the distribution. The very meaning of the word "career" contains a notion of dynamic progress. Yet, the issue of dynamic returns is largely ignored in the empirical literature on labor supply, presumably because of the challenges to estimating welfare-relevant, long-run elasticities in the presence of such returns. We have many compelling estimates of labor supply responses to tax reform, but they generally capture only short-term effects. In this paper, we propose a way to estimate long-term elasticities in the presence of dynamic returns without having to rely on a parameterized structural model.

We provide three contributions. First, we develop a new model of earnings responses to taxes in the presence of dynamic returns. In this model, the returns to effort are delayed and mediated by job switches such as promotions within firms or movements between firms. We use the model to provide a set of predictions that can be taken to the data, and to characterize how job switchers can be used to uncover the true long-run elasticity. Second, we provide descriptive evidence on earnings and hours-worked patterns over the lifecycle, verifying the predictions of the theoretical model. This analysis leverages the granularity and statistical power of the Danish data to provide particularly clean evidence. Third, informed by the model and descriptive evidence, we conduct a quasi-experimental study of earnings elasticities using job switchers. A conventional estimation approach gives a modest earnings elasticity of about 0.2, whereas our job switcher approach gives an elasticity close to 0.5. We present several analyses that address robustness and threats to identification, all of which support our empirical approach. A key advantage of the approach is that it does not require a unique experiment; it can be implemented using tax reform experiments of the type commonly used in the literature (see Saez, Slemrod, and Giertz 2012).

While we argue that the long-run elasticity is larger than typically estimated, our analysis does not support the very large elasticities implied by some macro calibrations. It is possible that our estimate remains a lower bound because job movers, while revealing more than job stayers, still do not reveal the full long-run effect. What are the policy implications of an elasticity of 0.5 (our long-run estimate) as opposed to an elasticity of 0.2 (standard short-run estimate)? Consider the Laffer rate on top earners. This is determined by the classic formula $\tau = 1/\left(1 + \varepsilon a\right)$, where ε is the earnings elasticity and a is the Pareto parameter (Diamond 1998; Saez 2001). Based on a Pareto parameter of 1.5 — the relevant number for the US — an elasticity of 0.2 implies $\tau = 0.77$, while an elasticity of 0.5 implies $\tau = 0.57$. Therefore, the long-run elasticity we estimate has significant policy implications. The implications are even larger in countries with more compressed earnings distributions because their larger Pareto parameter magnifies the impact of the elasticity.

References

ABOWD, JOHN M., AND DAVID CARD (1987): "Intertemporal Labor Supply and Long-Term Employment Contracts," *The American Economic Review*, 77(1), 50–68. 6

——— (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–445. 6

ABOWD, JOHN M., FRANCIS KRAMARZ, AND DAVID N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67(2), 251–333. 2, 6, 71

ALTONJI, JOSEPH G., AND CHRISTINA H. PAXSON (1986): "Job Characteristics and Hours of Work," *Research in Labor Economics*, 8A, 1–55. 6

——— (1988): "Labor Supply Preferences, Hours Constraints, and Hours-Wage Trade-offs," *Journal of Labor Economics*, 6(2), 254–276. 6

ANAGOL, SANTOSH, ALLAN DAVIDS, BENJAMIN B. LOCKWOOD, AND TARUN RAMADORAI (2022): "Diffuse Bunching with Frictions: Theory and Estimation," Working Paper. 6

BEST, MICHAEL CARLOS, AND HENRIK KLEVEN (2013): "Optimal Income Taxation with Career Effects of Work Effort," Working Paper. 2

BLUNDELL, RICHARD, AND THOMAS MACURDY (1999): "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, chap. 27. Elsevier. 4, 18

BRONSON, MARY ANN, AND PETER SKOGMAN THOURSIE (2021): "The Wage Growth and Within-Firm Mobility of Men and Women: New Evidence and Theory," Working Paper. 3

CABALLERO, RICARDO J., AND EDUARDO M.R.A. ENGEL (1991): "Dynamic (S,s) Economies," *Econometrica*, 59(6), 1659–1686. 8

CALVO, GUILLERMO A. (1983): "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12(3), 383–398. 8

CAPLIN, ANDREW, AND JOHN LEAHY (1991): "State-Dependent Pricing and the Dynamics of Money and Output," *The Quarterly Journal of Economics*, 106(3), 683–708. 8

CAPLIN, ANDREW S., AND DANIEL F. SPULBER (1987): “Menu Costs and the Neutrality of Money,” *The Quarterly Journal of Economics*, 102(4), 703–726. 8

CARD, DAVID, ANA RUTE CARDOSO, JOERG HEINING, AND PATRICK KLINE (2018): “Firms and Labor Market Inequality: Evidence and Some Theory,” *Journal of Labor Economics*, 36(S1), S13–S70. 2, 6

CARD, DAVID, ANA RUTE CARDOSO, AND PATRICK KLINE (2016): “Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women,” *The Quarterly Journal of Economics*, 131(2), 633–686. 2, 6

CARD, DAVID, JÖRG HEINING, AND PATRICK KLINE (2013): “Workplace Heterogeneity and the Rise of West German Wage Inequality,” *The Quarterly Journal of Economics*, 128(3), 967–1015. 2, 6

CHETTY, RAJ (2012): “Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply,” *Econometrica*, 80(3), 969–1018. 5

CHETTY, RAJ, JOHN N. FRIEDMAN, TORE OLSEN, AND LUIGI PISTAFERRI (2011): “Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records,” *The Quarterly Journal of Economics*, 126(2), 749–804. 6

CHETTY, RAJ, JOHN N. FRIEDMAN, AND EMMANUEL SAEZ (2013): “Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings,” *American Economic Review*, 103(7), 2683–2721. 6

CHETTY, RAJ, ADAM GUREN, DAY MANOLI, AND ANDREA WEBER (2013): “Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities,” *NBER Macroeconomics Annual*, 27(1), 1–56. 5

DIAMOND, PETER A. (1998): “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates,” *The American Economic Review*, 88(1), 83–95. 1, 7, 35

FARBER, HENRY S. (1999): “Mobility and Stability: The Dynamics of Job Change in Labor Markets,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3B, chap. 37. Elsevier. 6, 18

GRIGSBY, JOHN, ERIK HURST, AND AHU YILDIRMAZ (2021): “Aggregate Nominal Wage Adjustments: New Evidence from Administrative Payroll Data,” *American Economic Review*, 111(2), 428–471. 2

GRUBER, JON, AND EMMANUEL SAEZ (2002): "The Elasticity of Taxable Income: Evidence and Implications," *Journal of Public Economics*, 84(1), 1–32. 25, 30

GUVENEN, FATIH, FATIH KARAHAN, SERDAR OZKAN, AND JAE SONG (2021): "What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?," *Econometrica*, 89(5), 2303–2339. 2

HARRIS, MILTON, AND BENGT HOLMSTRÖM (1982): "A Theory of Wage Dynamics," *The Review of Economic Studies*, 49(3), 315–333. 6

HENDREN, NATHANIEL, AND BEN SPRUNG-KEYSER (2020): "A Unified Welfare Analysis of Government Policies," *The Quarterly Journal of Economics*, 135(3), 1209–1318. 1

HOLMSTRÖM, BENGT (1999): "Managerial Incentive Problems: A Dynamic Perspective," *The Review of Economic Studies*, 66(1), 169–182. 6

HUMLUM, ANDERS (2021): "Robot Adoption and Labor Market Dynamics," Working Paper. 16

IMAI, SUSUMU, AND MICHAEL P. KEANE (2004): "Intertemporal Labor Supply and Human Capital Accumulation," *International Economic Review*, 45(2), 601–641. 5

JACOBSON, LOUIS S., ROBERT J. LALONDE, AND DANIEL G. SULLIVAN (1993): "Earnings Losses of Displaced Workers," *The American Economic Review*, 83(4), 685–709. 4, 33

JAKOBSEN, KATRINE MARIE, AND JAKOB EGHOLT SØGAARD (2022): "Identifying Behavioral Responses to Tax Reforms: New Insights and a New Approach," *Journal of Public Economics*, 212. 5, 22, 24, 34

KEANE, MICHAEL, AND RICHARD ROGERSON (2015): "Reconciling Micro and Macro Labor Supply Elasticities: A Structural Perspective," *Annual Review of Economics*, 7(1), 89–117. 1, 2, 5

KEANE, MICHAEL P. (2011): "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 49(4), 961–1075. 2, 5

KLEVEN, HENRIK (2016): "Bunching," *Annual Review of Economics*, 8, 435–464. 6

——— (2021): "Sufficient Statistics Revisited," *Annual Review of Economics*, 13, 515–538. 10

——— (2024): "The EITC and the Extensive Margin: A Reappraisal," *Journal of Public Economics*, 236, 1–28. 5

Kleven, Henrik, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez (2011): "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark," *Econometrica*, 79, 651–692. 16

Kleven, Henrik, and Claus Thustrup Kreiner (2006): "The Marginal Cost of Public Funds: Hours of Work Versus Labor Force Participation," *Journal of Public Economics*, 90(10–11), 1955–1973. 1

Kleven, Henrik, Claus Thustrup Kreiner, and Emmanuel Saez (2009): "The Optimal Income Taxation of Couples," *Econometrica*, 77(2), 537–560. 7

Kleven, Henrik, and Mazhar Waseem (2013): "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *The Quarterly Journal of Economics*, 128(2), 669–723. 6

Kreiner, Claus Thustrup, Søren Leth-Petersen, and Peer Ebbesen Skov (2016): "Tax Reforms and Intertemporal Shifting of Wage Income: Evidence from Danish Monthly Payroll Records," *American Economic Journal: Economic Policy*, 8(3), 233–257. 5

Kreiner, Claus Thustrup, Jakob Roland Munch, and Hans Jørgen Whitta-Jacobsen (2015): "Taxation and the Long-Run Allocation of Labor: Theory and Danish Evidence," *Journal of Public Economics*, 127, 74–86. 6

Labanca, Claudio, and Dario Pozzoli (2022): "Constraints on Hours within the Firm," *Journal of Labor Economics*, 40(2), 473–503. 5, 6

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A. Woodbury (2023): "Work Hours Mismatch," Working Paper. 6

Nakamura, Emi, and Jón Steinsson (2010): "Monetary Non-Neutrality in a Multisector Menu Cost Model," *The Quarterly Journal of Economics*, 125(3), 961–1013. 8

Pavan, Ronni (2011): "Career Choice and Wage Growth," *Journal of Labor Economics*, 29(3), 549–587. 2, 6

Prescott, Edward C. (2004): "Why Do Americans Work So Much More Than Europeans?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1), 2–13. 1

ROGERSON, RICHARD, AND JOHANNA WALLENIUS (2009): "Micro and macro elasticities in a life cycle model with taxes," *Journal of Economic Theory*, 144(6), 2277–2292. 1

SAEZ, EMMANUEL (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *The Review of Economic Studies*, 68(1), 205–229. 1, 35

——— (2010): "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2(3), 180–212. 7

SAEZ, EMMANUEL, MANOS MATSAGANIS, AND PANOS TSAKLOGLOU (2012): "Earnings Determination and Taxes: Evidence From a Cohort-Based Payroll Tax Reform in Greece," *The Quarterly Journal of Economics*, 127(1), 493–533. 5

SAEZ, EMMANUEL, JOEL SLEMROD, AND SETH H. GIERTZ (2012): "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 50(1), 3–50. 4, 10, 22, 24, 35

SCHEUER, FLORIAN, AND IVÁN WERNING (2017): "The Taxation of Superstars," *The Quarterly Journal of Economics*, 132(1), 211–270. 5

SIGAARD, HANS SCHYTTE (2023): "Estimating labor supply responses to Danish tax reforms," *Journal of Public Economics*, 224, 104943. 5

TAYLOR, JOHN B. (1980): "Output and Price Stability: An International Comparison," *Journal of Economic Dynamics and Control*, 2, 109–132. 8

TOPEL, ROBERT H., AND MICHAEL P. WARD (1992): "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107(2), 439–479. 2, 6

TORTAROLO, DARIO, GUILLERMO CRUCES, AND VICTORIA CASTILLO (2020): "It takes two to tango: Labor responses to an income tax holiday in Argentina," Working Paper. 5

WILLIS, ROBERT J. (1986): "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and R. Layard, vol. 1, chap. 10. Elsevier. 18

TABLE 1: ESTIMATING THE EARNINGS ELASTICITY

	2-Year Elasticity			4-Year Elasticity		
	First Stage	Reduced Form	2SLS	First Stage	Reduced Form	2SLS
Panel A: Dynamic Approach (Switchers)						
$\Delta\log(1-\tau^{sim})$	0.486	0.236		0.441	0.215	
	(0.004)	(0.012)		(0.004)	(0.011)	
$\Delta\log(1-\tau)$			0.486			0.487
			(0.022)			(0.026)
Panel B: Standard Approach (All Workers)						
$\Delta\log(1-\tau)$			0.200			0.258
			(0.009)			(0.014)
N	2,501,126	2,501,126	2,501,126	2,428,411	2,428,411	2,428,411
R^2	0.480	0.169		0.456	0.156	
F-Stat	16,482.7			14,133.6		

Notes: This table presents estimates of the earnings elasticity with respect to the marginal net-of-tax rate. Panel A shows estimates for job switchers based on specification (25), while Panel B shows estimates for all workers regardless of job switching status. The first-stage and reduced-form estimations use simulated tax changes $\Delta\log(1-\tau^{sim})$, holding individual income fixed at its baseline level. The 2SLS estimations use actual tax changes $\Delta\log(1-\tau)$ instrumented with the simulated tax changes $\Delta\log(1-\tau^{sim})$. This gives the structural earnings elasticity of interest. We control for baseline job cell by including firm$\times$occupation fixed effects, but the results are robust to adding or dropping controls as we show below. The elasticities for all workers (Panel B) are based on scaling the switcher elasticities (Panel A) using the cumulated fraction of workers who have made a post-reform job switch over the given time horizon (as characterized by equation 7). Robust standard errors are shown in parentheses.

TABLE 2: ROBUSTNESS OF ESTIMATES

EFFECTS OF THE 2009 TAX REFORM UNDER ALTERNATIVE SPECIFICATIONS

	Earnings Elasticity		**Probability of Switching**	
	2-Year Effect	4-Year Effect	2-Year Effect	4-Year Effect
Panel A: Baseline Specification				
All Switchers	0.486 (0.022)	0.487 (0.026)	0.036 (0.001)	0.016 (0.001)
Top 10% Switchers	0.725 (0.029)	0.642 (0.029)	0.073 (0.002)	0.045 (0.002)
Panel B: Robustness (All Switchers)				
Alternative Controls				
No Controls	0.610 (0.020)	0.590 (0.024)	0.031 (0.001)	0.027 (0.001)
Demographics	0.582 (0.020)	0.524 (0.023)	0.029 (0.001)	0.024 (0.001)
Demographics & Job FE	0.465 (0.022)	0.437 (0.025)	0.037 (0.001)	0.018 (0.001)
Donut Specifications				
Small Donut	0.476 (0.023)	0.501 (0.026)	0.040 (0.001)	0.019 (0.001)
Large Donut	0.456 (0.023)	0.476 (0.027)	0.045 (0.001)	0.023 (0.002)
Type of Switch				
Occupation Switches	0.531 (0.024)	0.553 (0.028)	0.034 (0.001)	0.015 (0.001)
Firm Switches	0.491 (0.026)	0.504 (0.030)	0.004 (0.001)	-0.007 (0.001)
Mass-Layoff Switches	0.666 (0.031)	0.712 (0.049)	0.030 (0.001)	0.004 (0.001)

Notes: This table investigates the robustness of our estimates to alternative samples and specifications. The left columns show estimates of the earnings elasticity for job switchers, while the right columns show estimates of the effect on the probability of job switching. Our baseline specification controls for job cell fixed effects, includes all taxpayers with baseline income between 250,000-1,000,000 DKK, and uses all types of job switches in the estimation. Panel A explores heterogeneity by income level, comparing the effects for all treated switchers to the effects for treated switchers in the top decile. Panel B changes the specification by (i) dropping or adding controls, (ii) allowing for donut-hole specifications that exclude observations near the treatment threshold (a small donut of +/- 2.5% of the threshold or a large donut of +/- 5.0% of the threshold), and (iii) restricting attention to specific types of job switches (occupation, firm, or mass-layoff switches). The demographic controls include dummies for age, gender, marital status, children, occupational rank, and firm size. Mass-layoff switches are defined as firm switches in which the original firm reduced its workforce by at least 30% and had a workforce of least 20 employees.

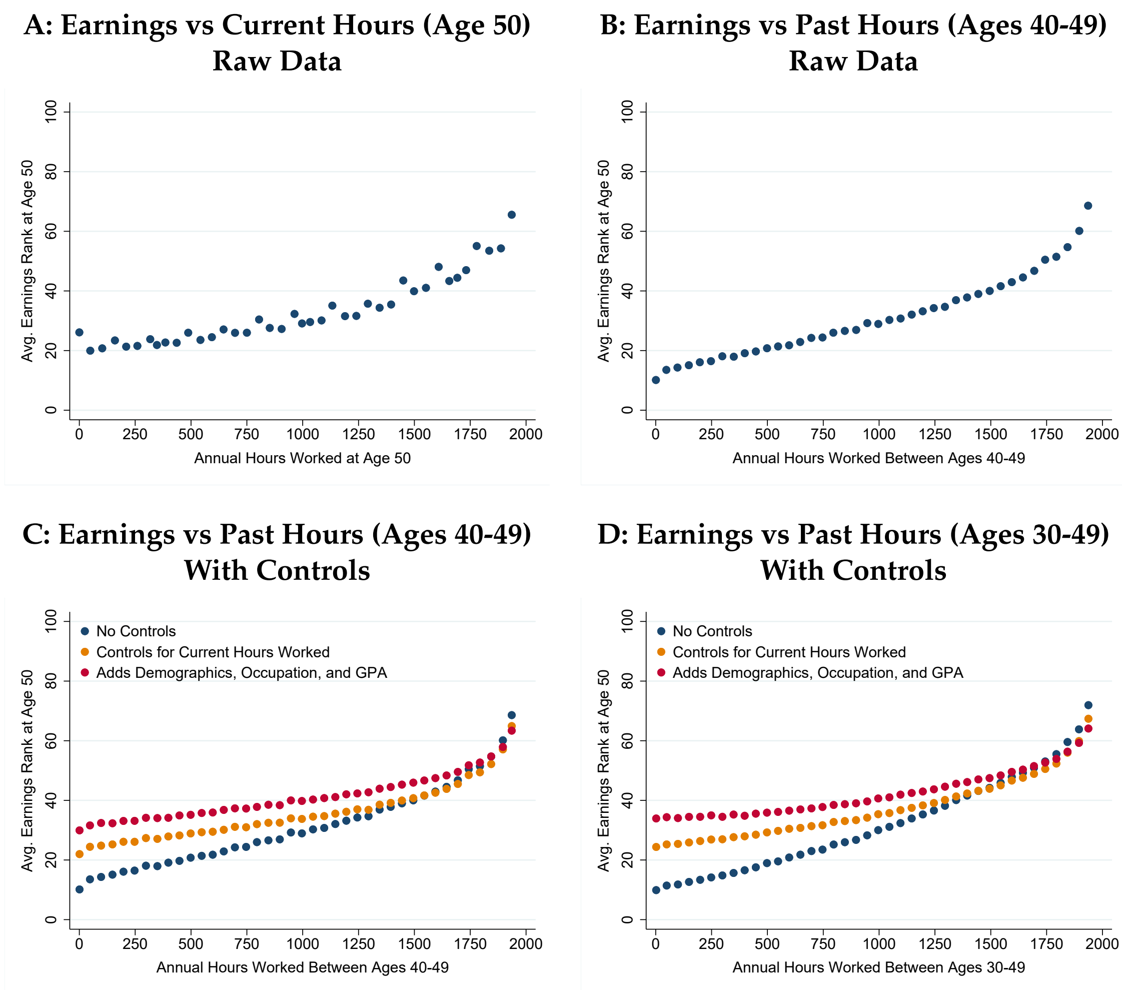
TABLE 3: EXTERNAL VALIDITY OF ESTIMATES

EFFECTS OF THE 1987 TAX REFORM UNDER ALTERNATIVE SPECIFICATIONS

	Earnings Elasticity		**Probability of Switching**	
	2-Year Effect	4-Year Effect	2-Year Effect	4-Year Effect
Panel A: Baseline Specification				
All Switchers	0.477 (0.054)	0.421 (0.044)	0.014 (0.002)	0.018 (0.002)
Top 10% Switchers	0.672 (0.070)	0.358 (0.052)	0.033 (0.002)	0.044 (0.003)
Panel B: Robustness (All Switchers)				
Alternative Controls				
No Controls	0.529 (0.052)	0.391 (0.042)	0.015 (0.002)	0.029 (0.002)
Demographics	0.555 (0.053)	0.419 (0.042)	0.004 (0.002)	0.011 (0.002)
Demographics & Job FE	0.471 (0.053)	0.424 (0.044)	0.010 (0.002)	0.016 (0.002)
Donut Specifications				
Small Donut	0.441 (0.055)	0.406 (0.046)	0.014 (0.002)	0.018 (0.002)
Large Donut	0.387 (0.058)	0.397 (0.050)	0.017 (0.002)	0.020 (0.002)
Type of Switch				
Occupation Switches	0.346 (0.059)	0.266 (0.048)	-0.000 (0.001)	-0.001 (0.002)
Firm Switches	0.609 (0.064)	0.548 (0.052)	0.017 (0.002)	0.015 (0.002)
Mass-Layoff Switches	0.718 (0.082)	0.534 (0.084)	0.003 (0.001)	0.002 (0.001)

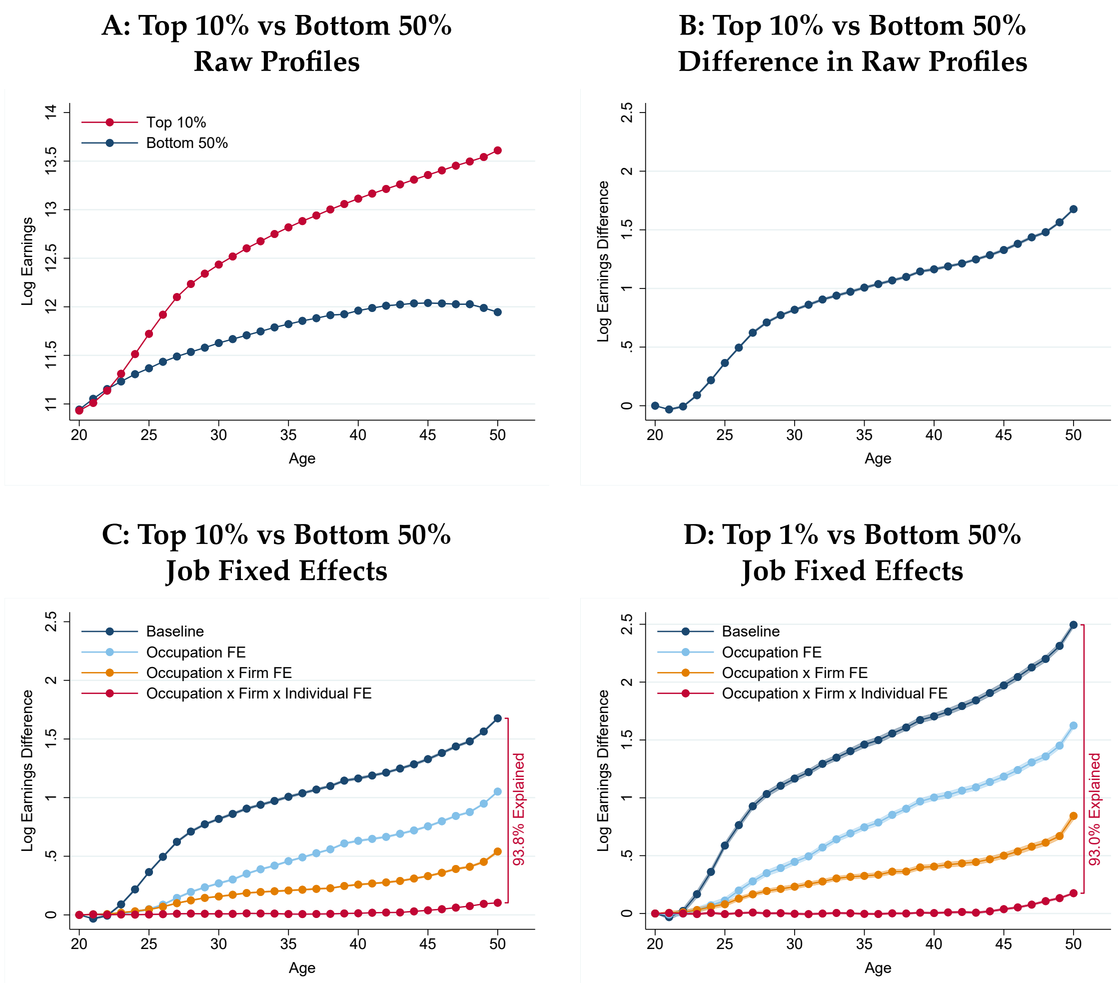
Notes: This table investigates the external validity of our estimates by using the 1987 tax reform for identification. Like the 2009 reform analyzed so far, the 1987 reform reduced tax rates at the top of the distribution relative to the tax rates further down. The table is constructed in the exact same was as the preceding tables for the 2009 reform.

FIGURE 1: PAST HOURS WORKED PREDICT EARNINGS, CONDITIONAL ON CURRENT HOURS



Notes: This figure shows that past hours worked predict current earnings. The figure is based on a balanced panel of workers observed between the ages of 20 and 50. Each panel plots the relationship between average earnings rank at age 50 and hours worked at different ages: current hours in Panel A and past hours in Panels B-D. The top panels depict raw data, while the bottom panels add controls. The predictive power of past hours worked remains strong even after controlling for current hours worked, demographic characteristics (dummies for gender, children, marital status, and education level), occupation (dummies for 2-digit DISCO codes), and school grades. School grades are measured as the GPA for the highest education degree obtained by age 21 (typically high school). The graphs include 95% confidence intervals based on standard errors clustered at the individual level, but these are hardly visible.

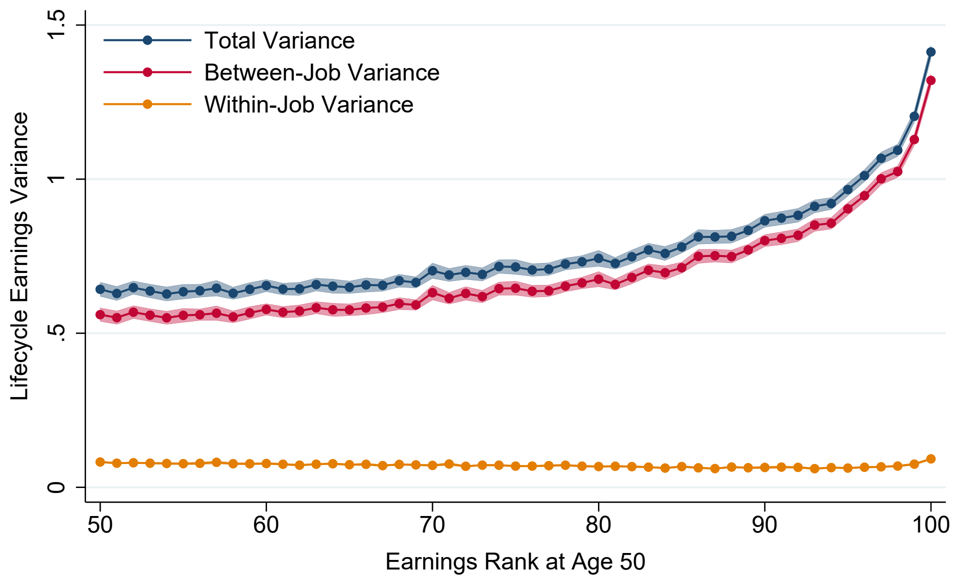
FIGURE 2: LIFECYCLE PROFILES IN EARNINGS ARE DRIVEN BY JOB SWITCHES



Notes: This figure shows that the lifecycle profile of earnings for workers who reach the top of the distribution is driven by job switches, defined as transitions between occupation×firm cells. The figure is based on a balanced panel of workers observed between the ages of 20 and 50, plotting the earnings profiles of workers observed in different parts of the distribution at age 50. Panel A plots the raw profiles for workers in the top 10% and bottom 50%, respectively. Panel B plots the difference in these raw profiles. Panel C compares the difference in raw profiles (dark blue) to the differences net of occupation fixed effects (light blue), net of occupation×firm fixed effects (orange), and net of occupation×firm×individual fixed effects (red). Within individual, job fixed effects explain about 94% of the lifecycle divergence between top-10% and bottom-50% earners. Panel D repeats the analysis of Panel C, but for top-1% earners. The shaded areas (not always visible) represent 95% confidence intervals based on standard errors clustered at the individual level.

FIGURE 3: BETWEEN-JOB VS WITHIN-JOB VARIATION IN EARNINGS

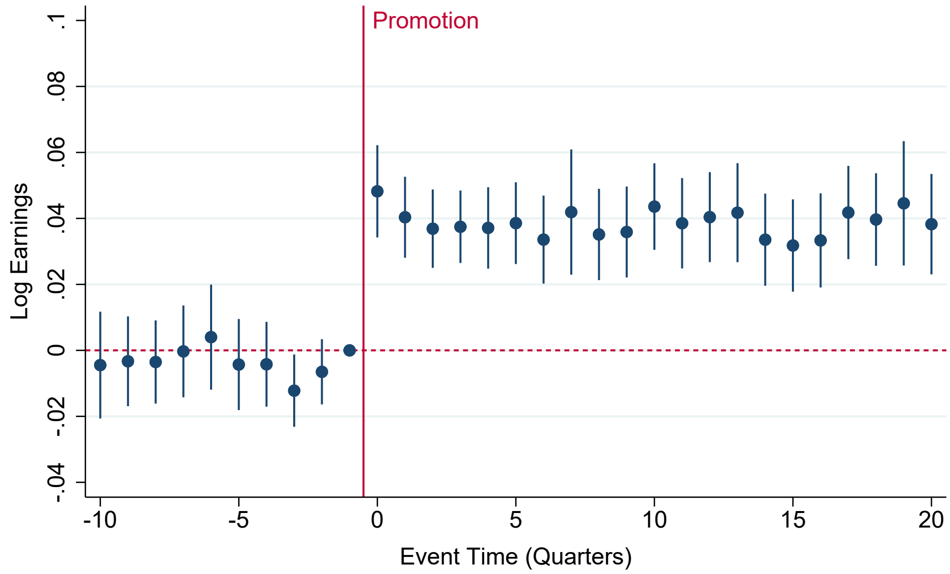
VARIANCE DECOMPOSITION BY EARNINGS RANK AT AGE 50



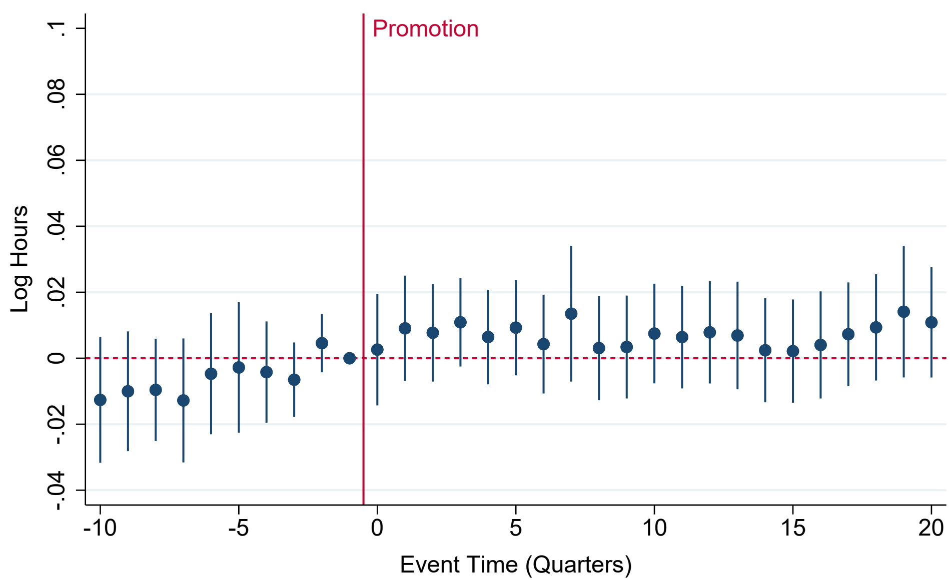
Notes: This figure decomposes the variance of lifecycle earnings into between-job variance and within-job variance using equation (22). Jobs are measured as occupation×firm cells and the sample is a balanced panel of workers observed between the ages of 20 and 50. The figure plots the total variance of earnings (blue), the between-job variance (red), and the within-job variance (yellow) by earnings rank at age 50. At all ranks shown, almost all of the lifecycle variation in earnings can be attributed to between-job variation, i.e. to switches between occupation×firm cells. The shaded areas depict 95% confidence intervals based on standard errors clustered at the individual level.

FIGURE 4: EARNINGS JUMP DISCRETELY AT PROMOTIONS, WITH NO CHANGE IN HOURS

A: Event Study of Earnings

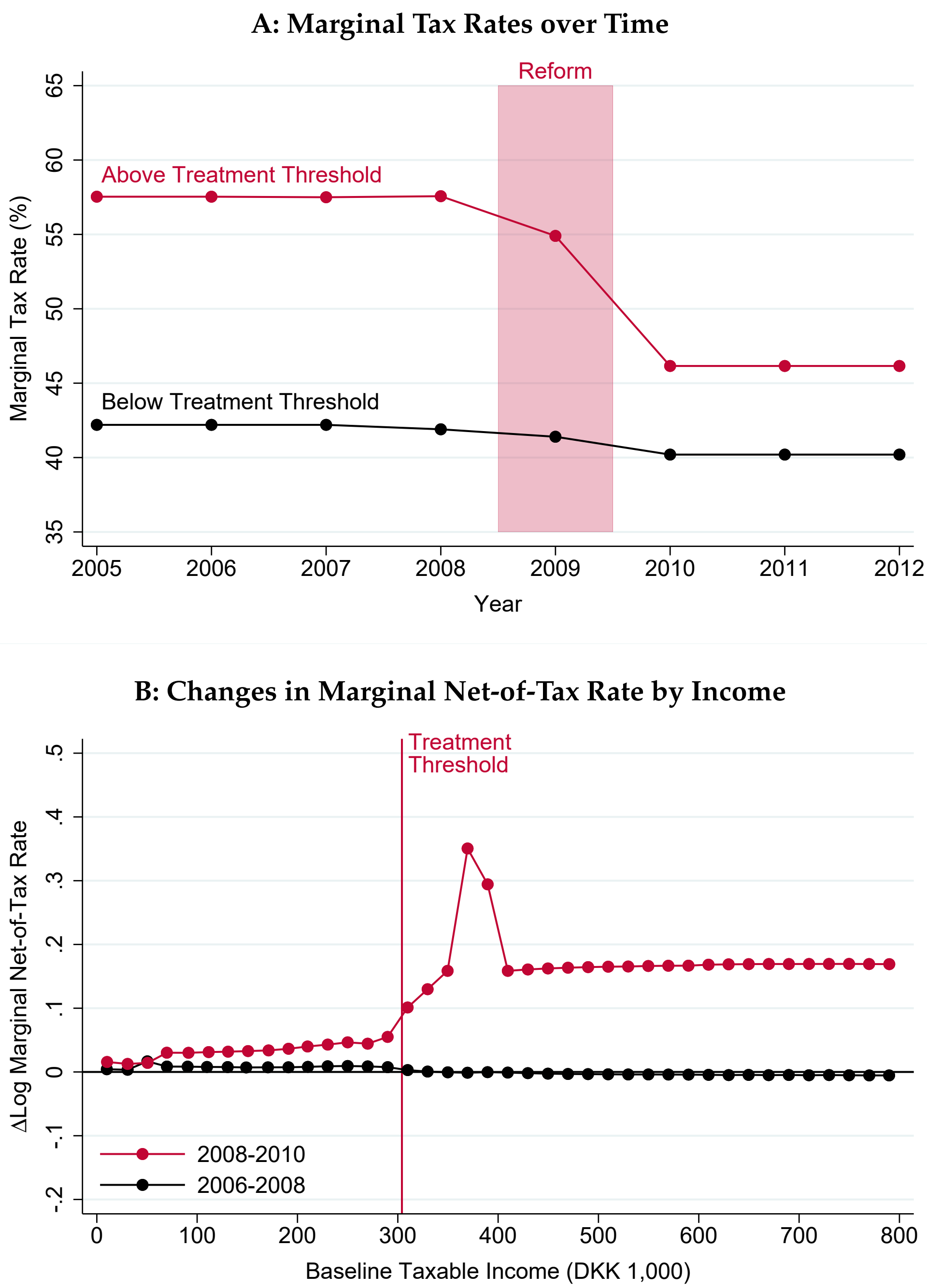


B: Event Study of Hours Worked



Notes: This figure presents event studies of promotions using specification (23). A promotion is defined as a switch to an occupation cell in which median earnings are at least 10% higher. The event study series show the outcomes of promoted workers relative to their unpromoted co-workers by quarter, normalizing the pre-promotion quarter to zero. Panel A considers earnings and Panel B considers hours worked. Promotions lead to sharp jumps in earnings, while hours worked are smooth. The error bars depict 95% confidence intervals based on standard errors clustered at the individual level.

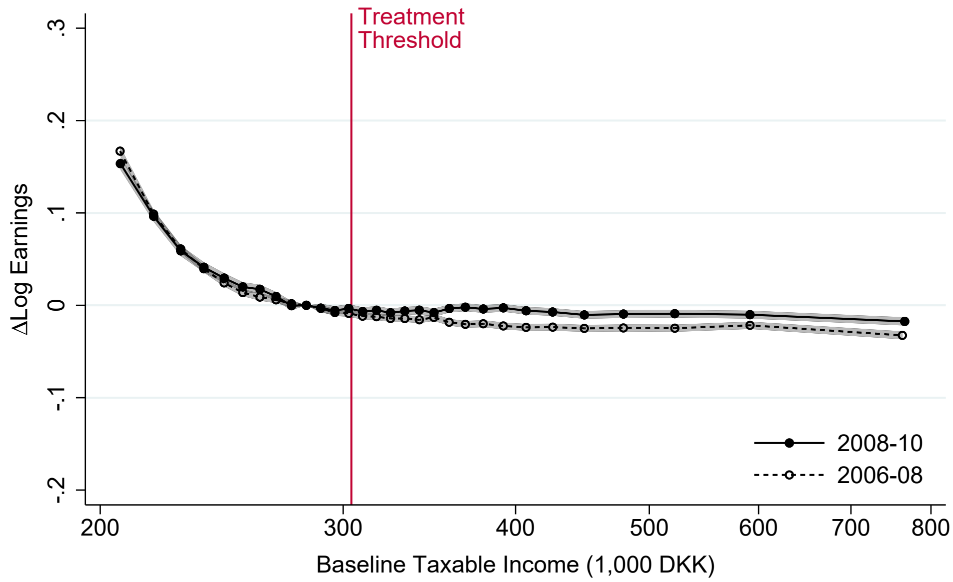
FIGURE 5: 2009 TAX REFORM IN DENMARK



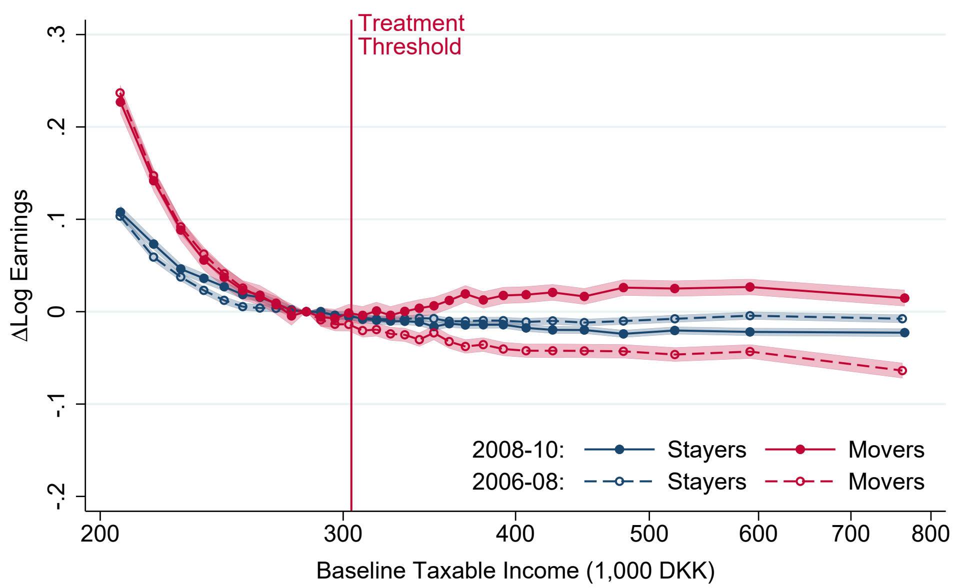
Notes: This figure illustrates the tax variation created by the 2009 reform in Denmark. Panel A shows the evolution of marginal tax rates for taxpayers above and below a treatment threshold, located at around the 70th percentile of the income distribution in 2008. The reform reduced marginal tax rates above the threshold by 11.4pp on average, while leaving marginal tax rates below the threshold virtually unchanged. Panel B shows that there was heterogeneity in the tax rate reduction depending on exact income location. This panel plots changes in the log marginal net-of-tax rate, $\Delta\log(1-\tau)$, by taxable income bin between 2008-10 (reform period) and 2006-08 (pre-reform period). The 2009 reform abolished the middle tax (increasing $1-\tau$ by about 17 log points above the treatment threshold) and raised the top tax threshold (increasing $1-\tau$ by another 20 log points around the 350K-400K range).

Figure 6: Impact of Tax Reform on Earnings

A: All Workers



B: Job Movers vs Job Stayers

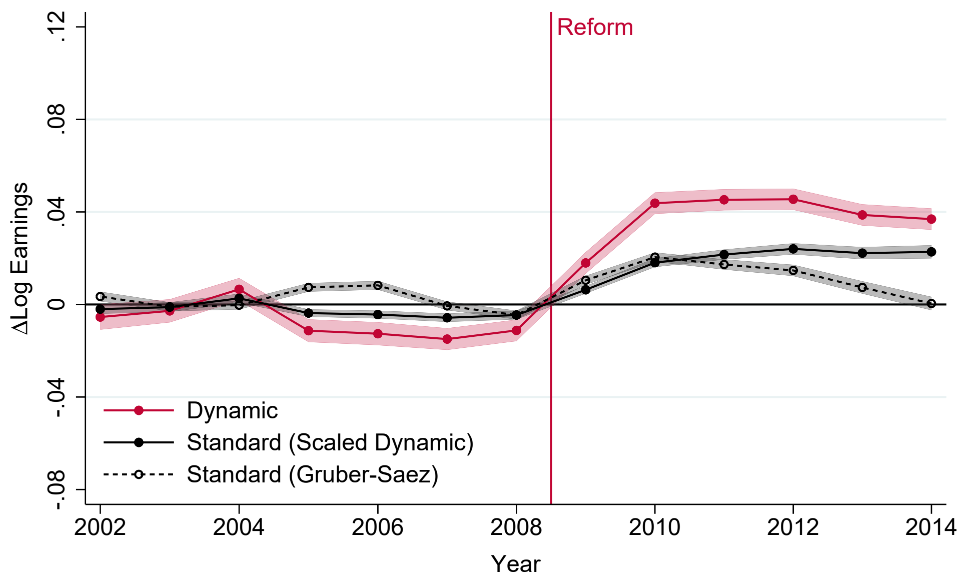


Notes: This figure shows the impact of the 2009 tax reform on earnings for all workers (Panel A) and for job movers vs job stayers (Panel B). Job movers are defined as those who switch between firm×occupation cells. Each panel plots changes in log earnings between 2008-10 (reform period) and 2006-08 (pre-reform, placebo period) by baseline taxable income. The 2006-08 series capture non-tax effects on earnings growth that vary by income level (including mean reversion). The 2008-10 and 2006-08 series track each other below the treatment threshold, but diverge above the threshold. This divergence represents the causal impact of the tax reform on earnings. Panel A shows a precisely estimated, but modest earnings increase in the full sample of treated workers. Panel B shows that the modest average impact masks strong heterogeneity by job switching status. Job movers feature large earnings increases, while job stayers feature no earnings increases. The shaded areas show 95% confidence intervals based on robust standard errors.

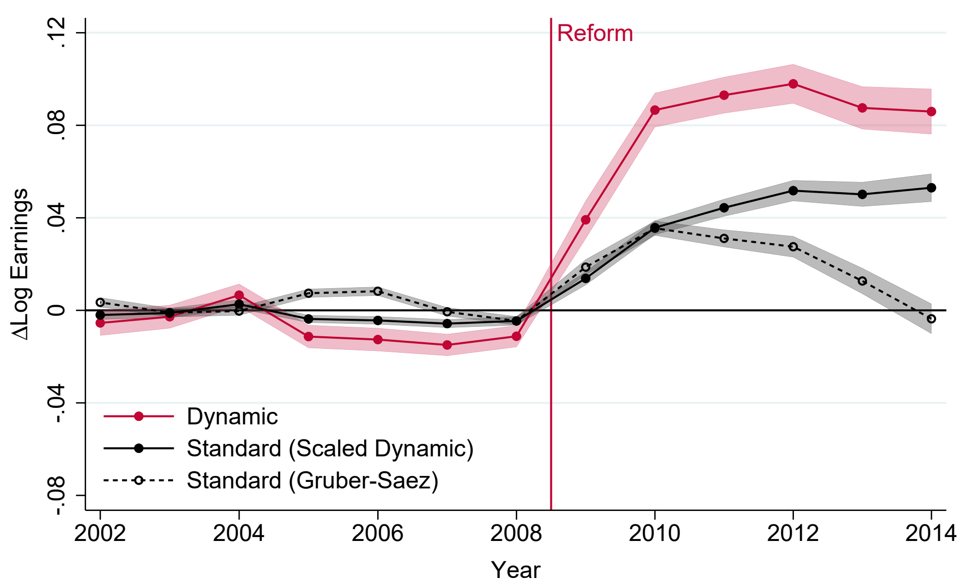
FIGURE 7: IMPACT OF TAX REFORM ON EARNINGS OVER TIME

DYNAMIC APPROACH (MOVERS) VS STANDARD APPROACH (ALL WORKERS)

A: Intention-to-Treat (ITT)

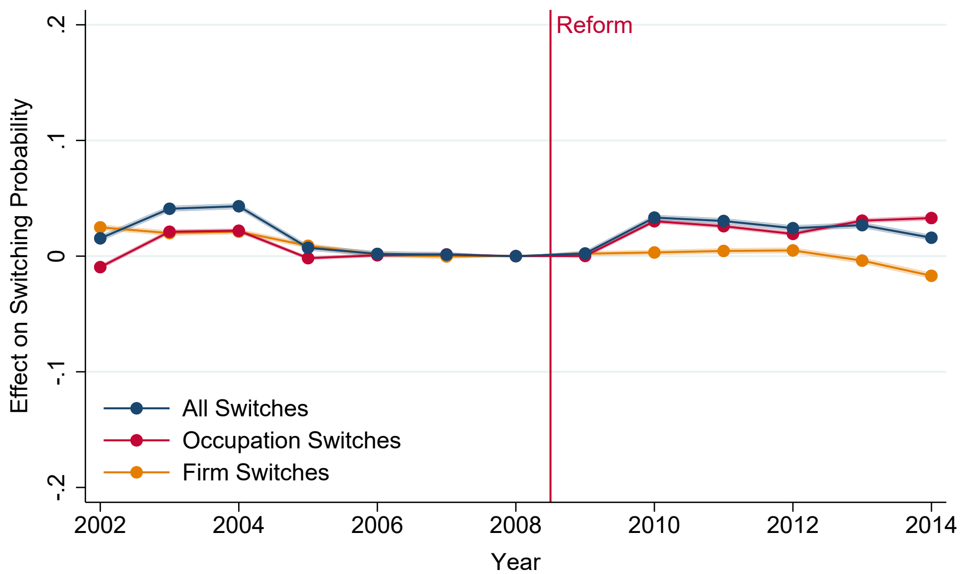


B: Treatment-on-the-Treated (TOT)



Notes: This figure shows the time profile of ITT effects (Panel A) and TOT effects (Panel B) on earnings. The red series show earnings responses for job movers (dynamic approach) based on the quadruple-differences approach described in section 5.1. The pre-reform part of the series (2002-2008) gives placebo estimates by comparing, for each year t, earnings growth between year $t-2$ and t to earnings growth between year $t-4$ and $t-2$. The post-reform part of the series (2009-2014) gives cumulative estimates by comparing, for each year t, earnings growth between 2008 and year t to earnings growth between 2006 and 2008. The mover estimates are compared to standard estimates of earnings responses in the full sample, including job stayers. One approach (solid black) is based on scaling the mover estimates by the cumulated fraction of workers who have made a post-reform job switch, thus imputing a zero earnings responses for stayers. The other approach (dashed black) estimates earnings responses based on a Gruber-Saez specification, avoiding the imputation of stayer responses to zero. This specification performs poorly over longer time horizons due to mean reversion effects. The figure demonstrates that (i) the dynamic mover estimates are stable over time (consistent with point identification of the long-run response) and (ii) the long-run response cannot be estimated from standard approaches by extending the time window.

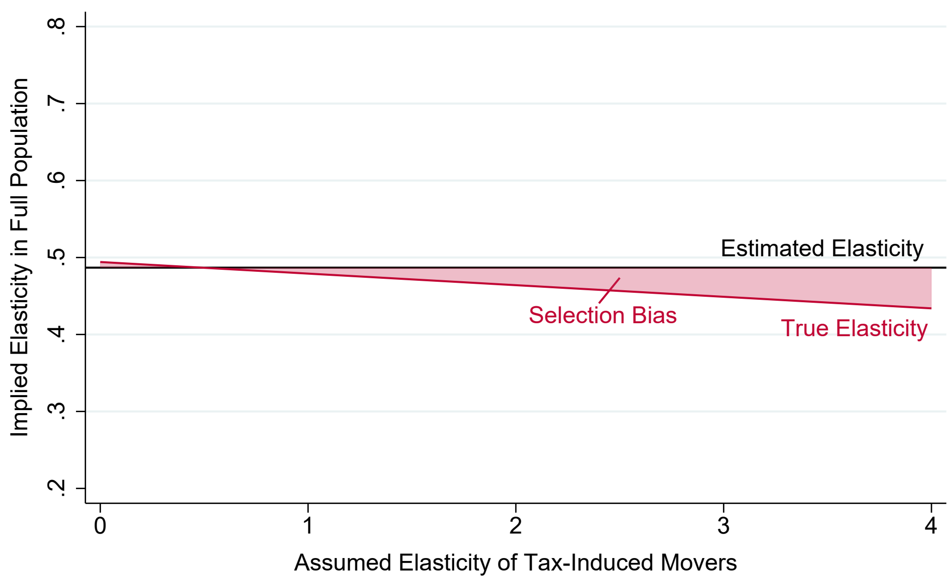
FIGURE 8: IMPACT OF TAX REFORM ON SWITCHING PROBABILITY



Notes: This figure shows the impact of the 2009 tax reform on the probability of switching jobs estimated based on the event study specification (26). The figure shows effects on all switches between occupations or firms (blue series), occupation switches (red series), and firm switches (yellow series). Up until 2010, we consider effects on the switching probability over 2-year intervals (2000-2002,...,2008-10). After 2010, we gradually increase the time interval to always include the entire post-reform period (2008-2011,...,2008-14). The shaded areas (hardly visible) show 95% confidence intervals based on robust standard errors.

Figure 9: Is Switching Selected?

Bounds on Selection Bias from Tax-Induced Switching

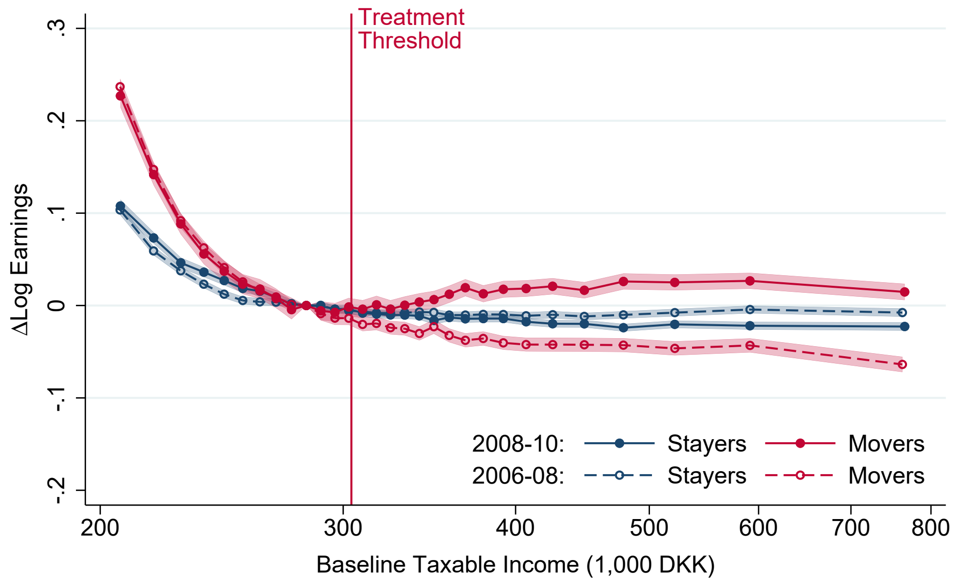


Notes: This figure provides bounds on the amount of selection bias due to tax-induced job switching. The analysis is based on the characterization in equations (27)-(28). It relies on (i) our estimate of the effect of the tax reform on the switching probability and (ii) assumptions about the structural elasticity for tax-induced movers (shown on the x-axis). The figure compares our baseline estimate of the earnings elasticity for job movers (black line) to the true long-run earnings elasticity (red line) as a function of the assumed elasticity for tax-induced movers. The amount of selection bias is modest even under extreme elasticities. The reason is that the effect of the reform on the switching probability is too small to create non-trivial selection bias.

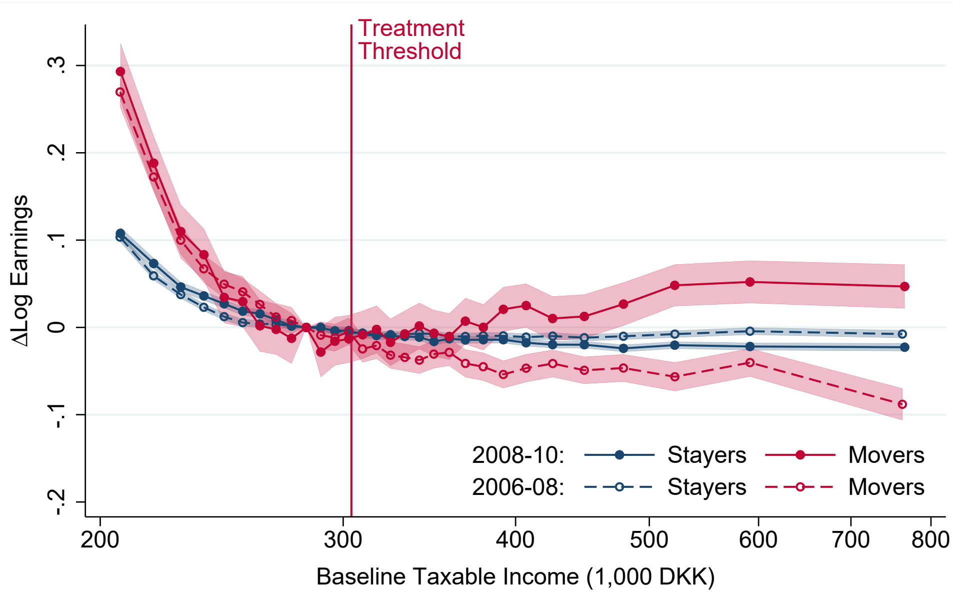
FIGURE 10: IS SWITCHING SELECTED?

ALL SWITCHERS VS MASS-LAYOFF SWITCHERS

A: All Switchers



B: Mass-Layoff Switchers



Notes: This figure shows the impact of the 2009 tax reform on earnings in the full sample of switchers (Panel A) and in a sample of mass-layoff switchers (Panel B). Switchers are those who move firm and/or occupation, while mass-layoff switchers are those who move firm following a mass layoff in their original firm. To qualify as a mass layoff, we require that a firm reduces its workforce by at least 30% on a base of at least 20 employees. Each panel plots changes in log earnings between 2008-10 (reform period) and between 2006-08 (pre-reform, placebo period) by baseline income bin. The earnings responses in the mass-layoff sample are similar to those in the full sample of switchers. This suggests that our switcher-based approach to estimating long-run earnings elasticities is not biased by selection into switching. The shaded areas show 95% confidence intervals based on robust standard errors.

Online Appendix

A Supplementary Tables and Figures

TABLE A.1: TOP VS BOTTOM OCCUPATIONS

JOB TITLES AND EARNINGS

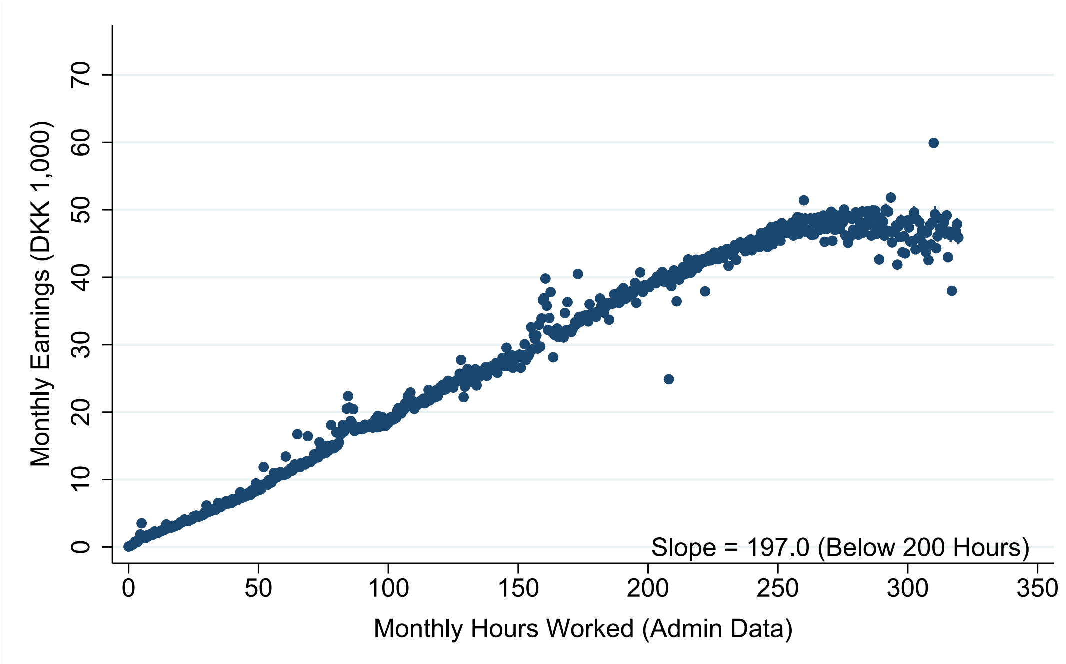
	Occupation	Earnings (DKK 1,000)		
		Mean	**P10**	**P90**
Top 10	Top Executives	2,107.6	447.7	5,181.2
	Managing Directors	1,071.7	407.8	1,890.2
	Securities and Currency Traders	1,050.7	397.5	1,963.2
	Administrative Directors	1,024.0	382.2	1,803.6
	Lawyers	983.6	401.5	1,810.8
	Pilots	929.7	476.7	1,400.5
	Medical Doctors	899.9	486.3	1,287.4
	Senior Government Officials	871.4	492.8	1,432.2
	Finance and Insurance Analysts	849.7	442.8	1,332.6
	Managers, Police and Judiciary	843.6	703.2	1,041.3
Bottom 10	Retail Assistants	268.8	56.8	484.0
	Machine Operators	268.8	59.6	527.3
	Cleaners	266.0	180.3	355.6
	Street and Market Sales Persons	260.3	29.6	481.6
	Services and Sales Workers	258.1	73.7	460.8
	Tailors	257.2	57.6	439.3
	Couriers	256.1	122.2	399.2
	Pottery Makers	240.4	67.1	400.6
	Beauticians	235.6	57.5	438.5
	Manual Laborers, Agriculture	221.4	103.0	343.3

Notes: This figure shows the highest-paying and lowest-paying occupations among workers aged 45-50. The classification is based on 6-digit occupation codes, ranked by mean wage earnings. For each occupation cell, the table reports the mean, the 10th percentile (P10), and the 90th percentile (P90) of wage earnings.

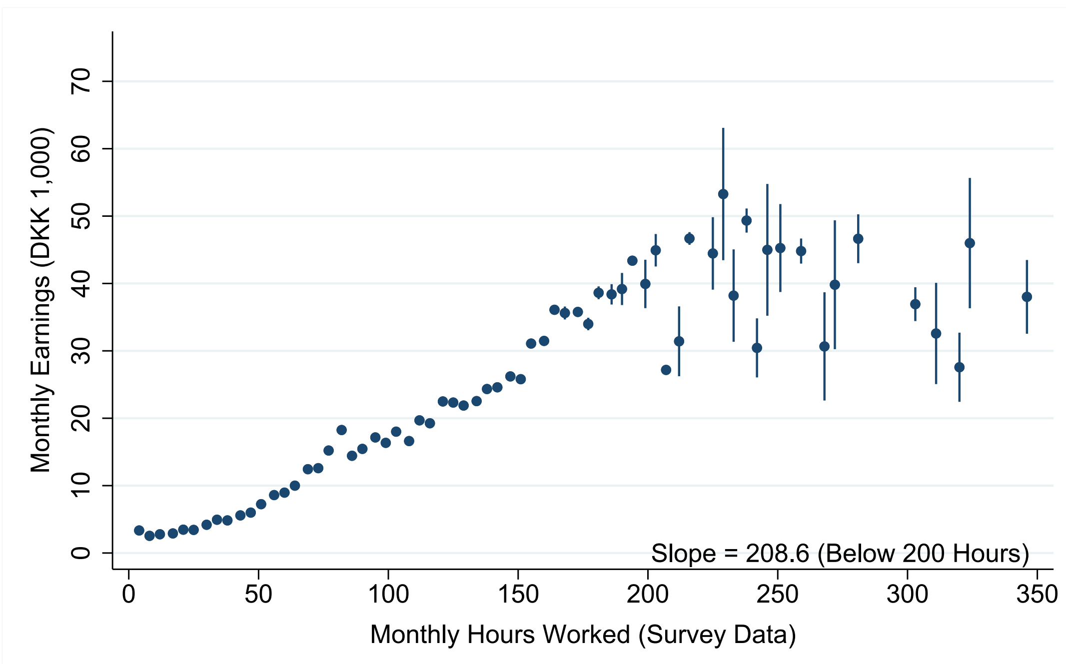
FIGURE A.1: VALIDATION OF ADMINISTRATIVE HOURS WORKED MEASURE

ADMINISTRATIVE DATA VS SURVEY DATA

A: Administrative Data (Pension Measure of Hours Worked)

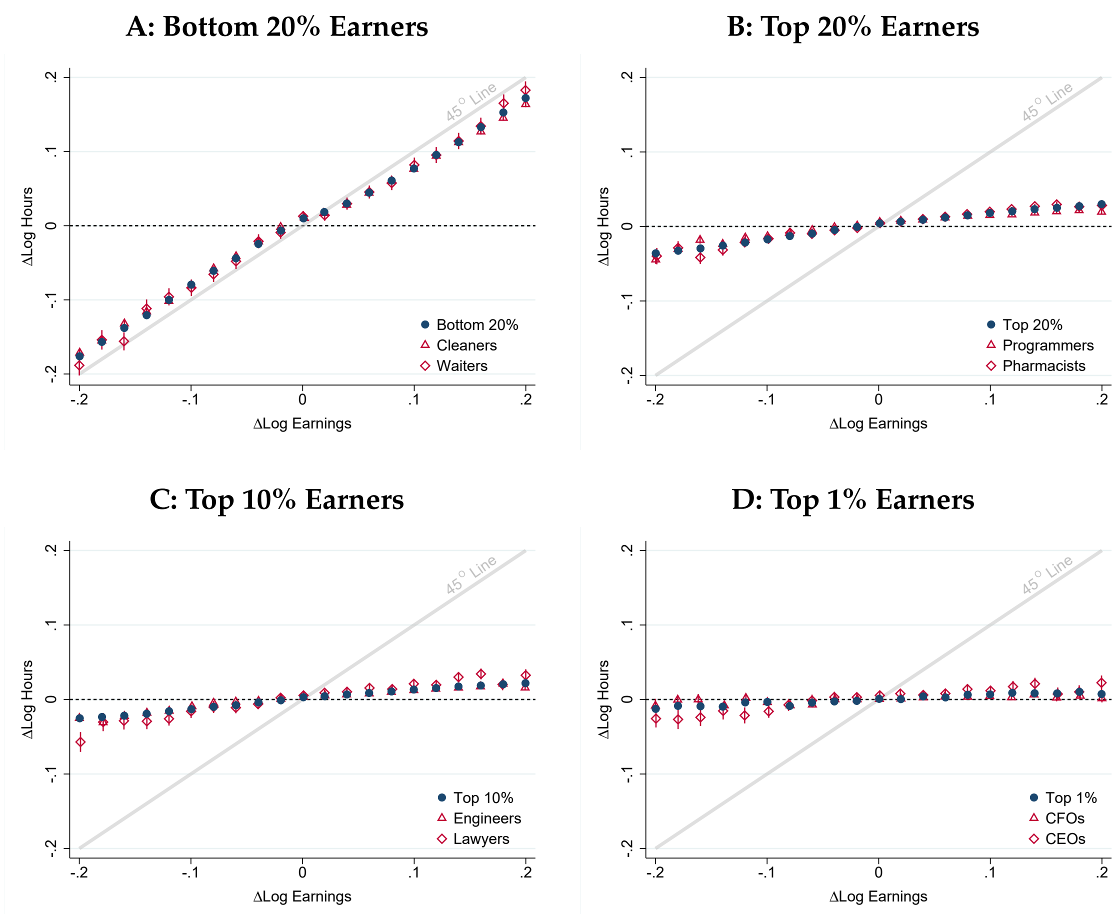


B: Survey Data (Self-Reported Measure of Hours Worked)



Notes: This figure validates the administrative measure of hours worked — the pension measure described in section 3 — against a survey measure of hours worked. The survey measure is based on a question about actual, uncapped hours taken from the Danish component of the EU Labour Force Survey. The figure plots the relationship between earnings and hours worked in the administrative data (Panel A) and in the survey data (Panel B). The earnings-hours relationship is similar in the two data sources. However, the survey measure is much more noisy than the administrative measure, especially at the top of the hours and earnings distribution, which is a key reason for using administrative data. The error bars depict 95% confidence intervals.

Figure A.2: Contemporaneous Hours and Earnings Changes are Unrelated at the Top, but not at the Bottom

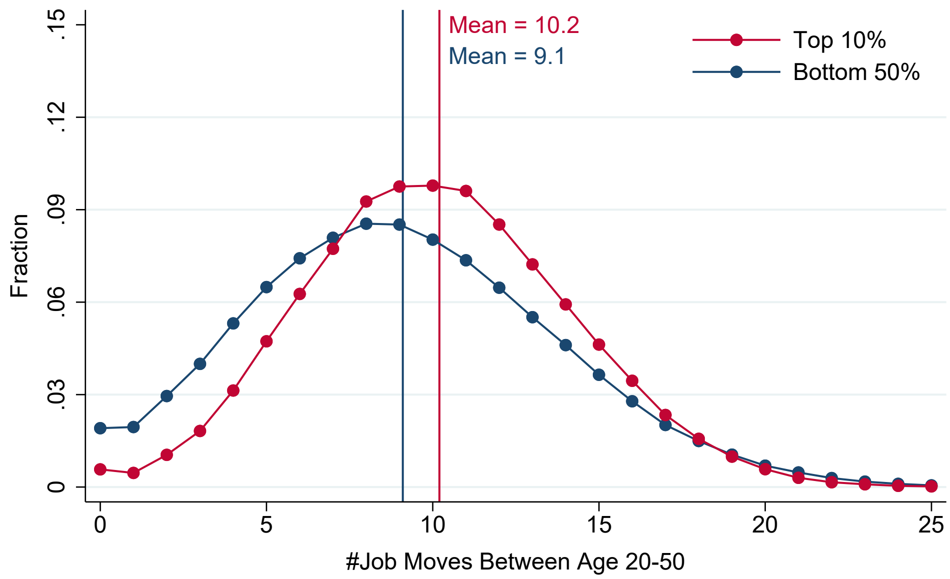


Notes: This figure shows the contemporaneous relationship between hours and earnings changes at the intensive margin. It plots changes in log hours against changes in log earnings in different segments of the earnings distribution: the bottom 20%, the top 20%, the top 10%, and the top 1%. The average relationship in each segment is depicted by blue dots, while examples of representative occupations in the different segments are depicted by red triangles and diamonds. While hours and earnings changes are almost perfectly correlated at the bottom (consistent with hourly-paid workers), they are virtually uncorrelated at the top (consistent with salaried workers). The error bars depict 95% confidence intervals based on standard errors clustered at the individual level.

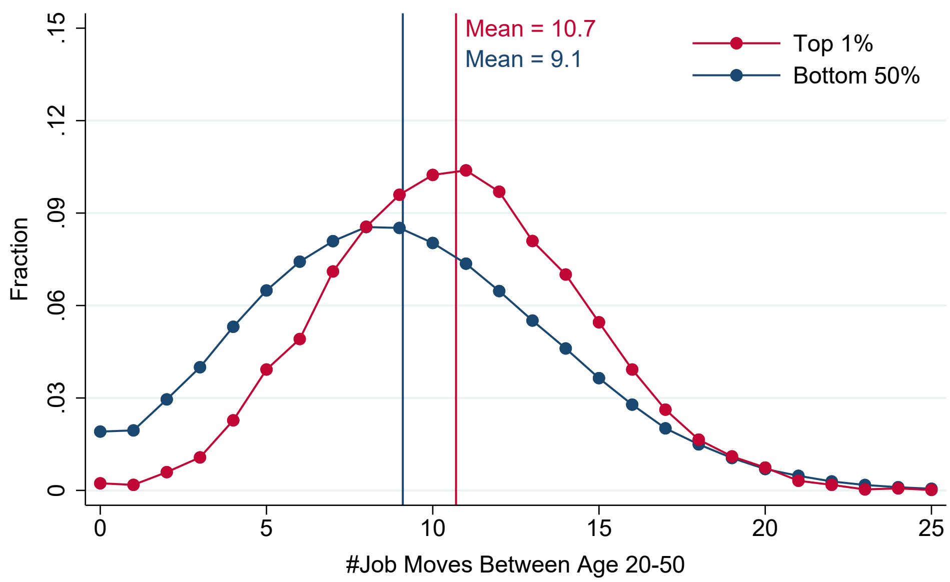
FIGURE A.3: DISTRIBUTION OF NUMBER OF SWITCHES

TOP VS BOTTOM EARNERS BETWEEN AGES 20-50

A: Top 10% vs Bottom 50%



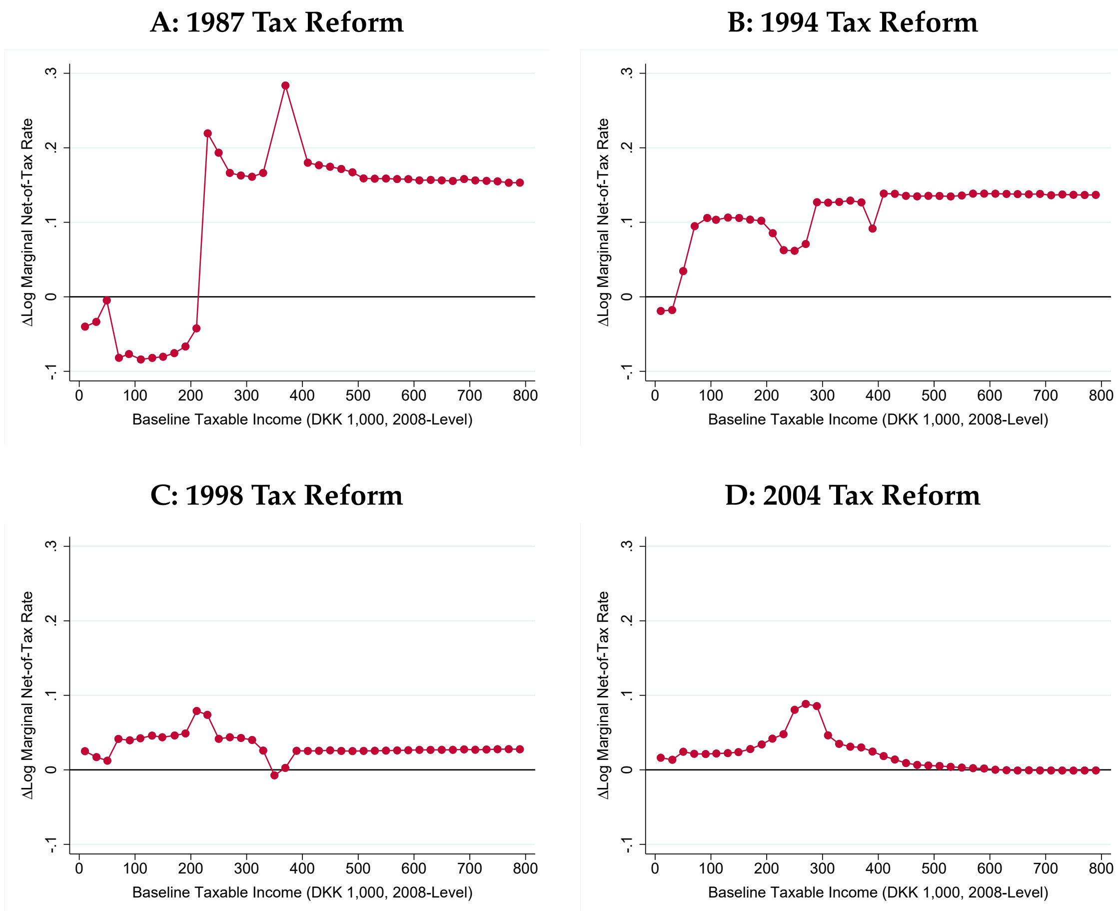
B: Top 1% vs Bottom 50%



Notes: This figure shows the distribution of the number of job switches between ages 20-50 for top and bottom earners. The figure is based on a balanced panel of workers observed between ages 20-50, splitting the sample by their earnings percentile at age 50. Panel A compares top-10% and bottom-50% earners, while Panel B compares top-1% and bottom-50% earners. The distributions are broadly similar for top and bottom earners. The average number of job switches is about 10 at the top and 9 at the bottom, corresponding to roughly one switch every three years.

FIGURE A.4: MAJOR TAX REFORMS PRIOR TO THE 2009 REFORM

REFORM-INDUCED CHANGES IN THE MARGINAL NET-OF-TAX RATE

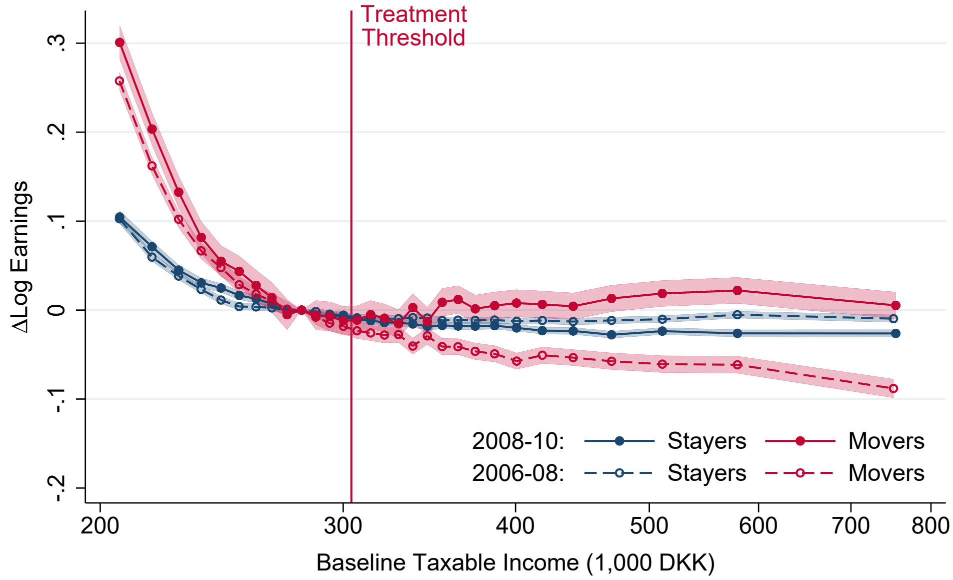


Notes: This figure illustrates the tax variation created by four major tax reforms implemented in Denmark prior to the 2009 reform. Each panel plots the change in the log marginal net-of-tax rate $1 - \tau$ by income bin. The general theme of Danish tax reforms since the 1980s has been to lower marginal tax rates while broadening the tax base. As can be seen from the figure, only the 1987 reform created the kind of tax variation needed for our analysis: tax changes on top earners relative to bottom earners. In fact, the 1987 reform is quite similar to the 2009 reform (shown in Figure 5) in terms of the magnitude and distribution of tax changes.

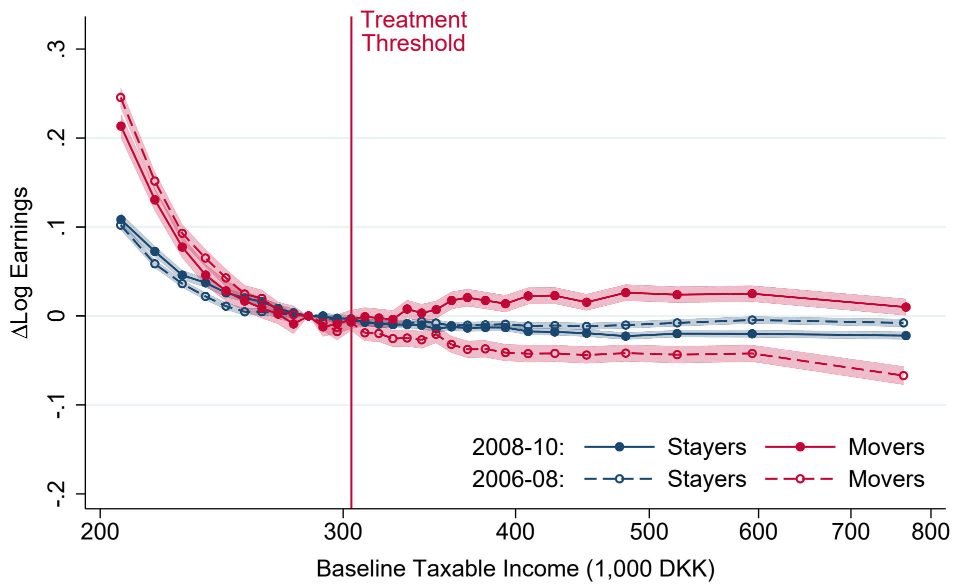
FIGURE A.5: IMPACT OF TAX REFORM ON EARNINGS

BY TYPE OF SWITCH

A: Firm Switches



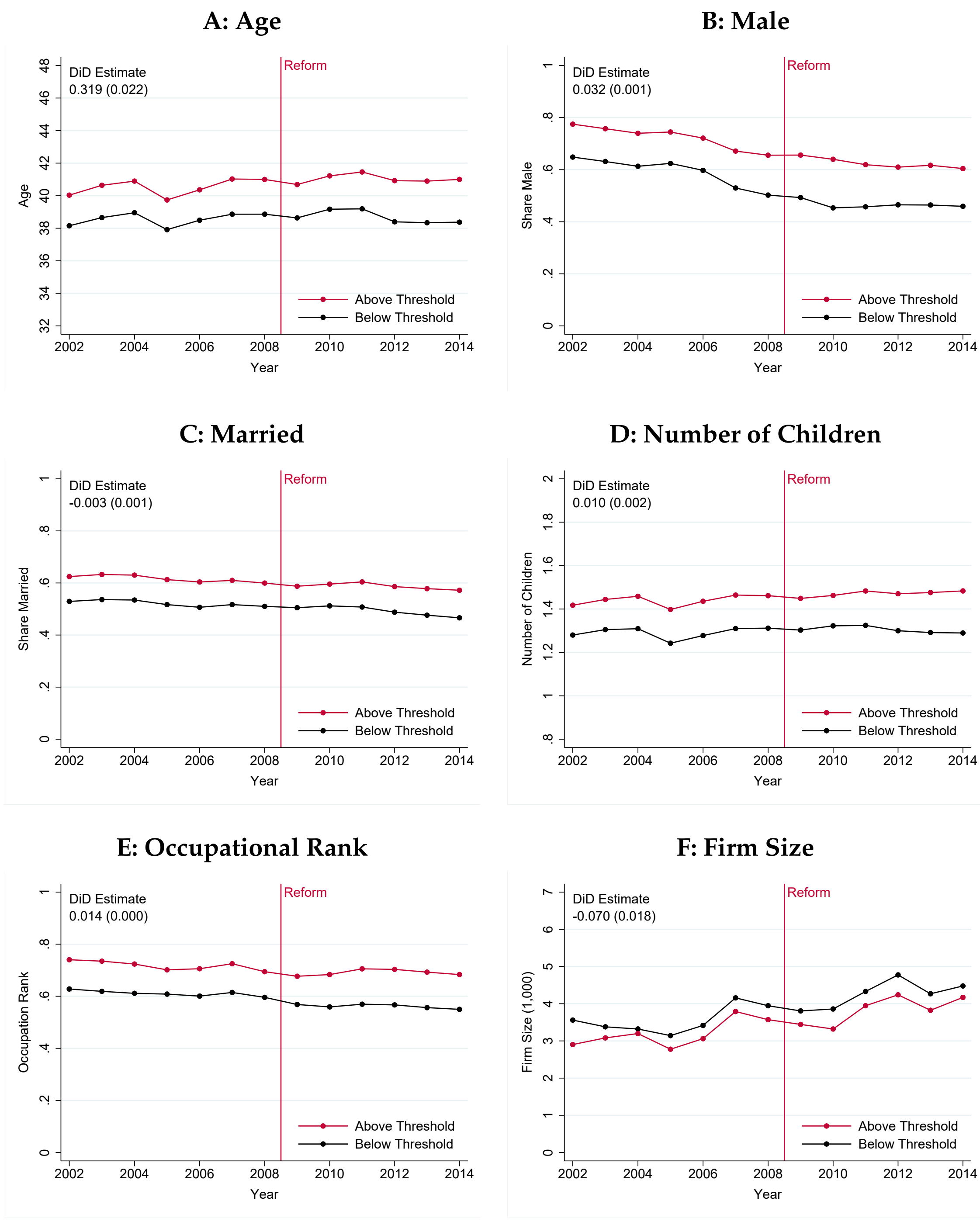
B: Occupation Switches



Notes: This figure shows the impact of the 2009 tax reform on earnings for firm switchers (Panel A) and occupation switchers (Panel B), each of them compared to non-switchers. To retain statistical power, Panel A includes all firm switchers (even if they also switch occupation) while Panel B includes all occupation switchers (even if they also switch firm). The figure is otherwise constructed in the same way as Figure 6. It plots changes in log earnings between 2008-10 (reform period) and between 2006-08 (pre-reform, placebo period) by baseline income bin. The empirical patterns are very similar in the two samples, with large earnings responses among both firm and occupation movers. The shaded areas show 95% confidence intervals based on robust standard errors.

FIGURE A.6: IS SWITCHING SELECTED?

IMPACT OF TAX REFORM ON SWITCHER CHARACTERISTICS

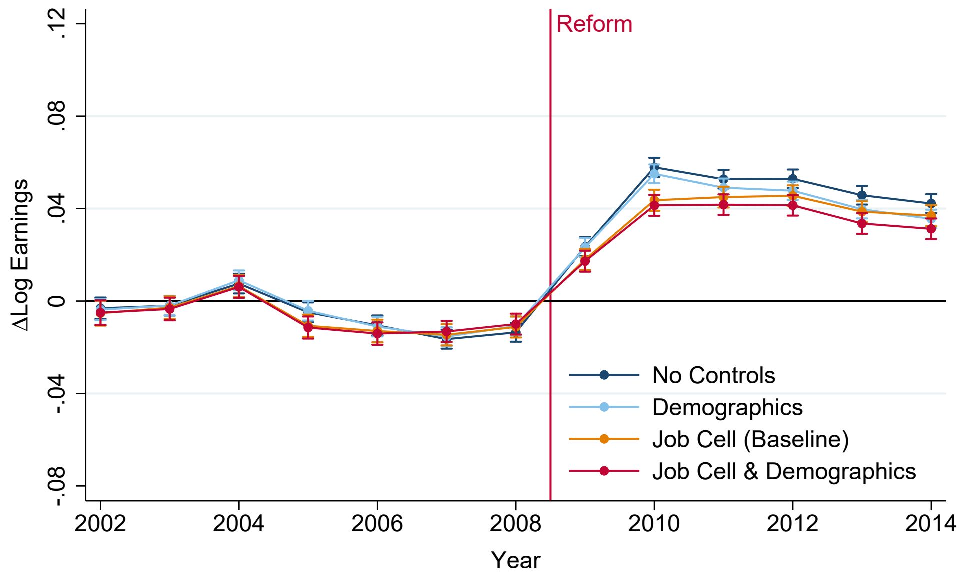


Notes: This figure investigates if the observable characteristics of treated job switchers relative to untreated job switchers diverge after the 2009 tax reform. Each panel plots the time series of a demographic variable for job switchers above and below the treatment threshold. Six different variables are considered: age, fraction male, fraction married, number of children, occupational rank, and firm size. The measure of occupational rank is based on ordering occupation cells by their mean earnings. The figure reports difference-in-differences estimates of the effect on each variable. These estimates are very small, albeit statistically significant due to the statistical power of our data. The analysis implies that job switching is not selected on observables given our quasi-experimental design. The shaded areas (hardly visible) show 95% confidence intervals based on robust standard errors.

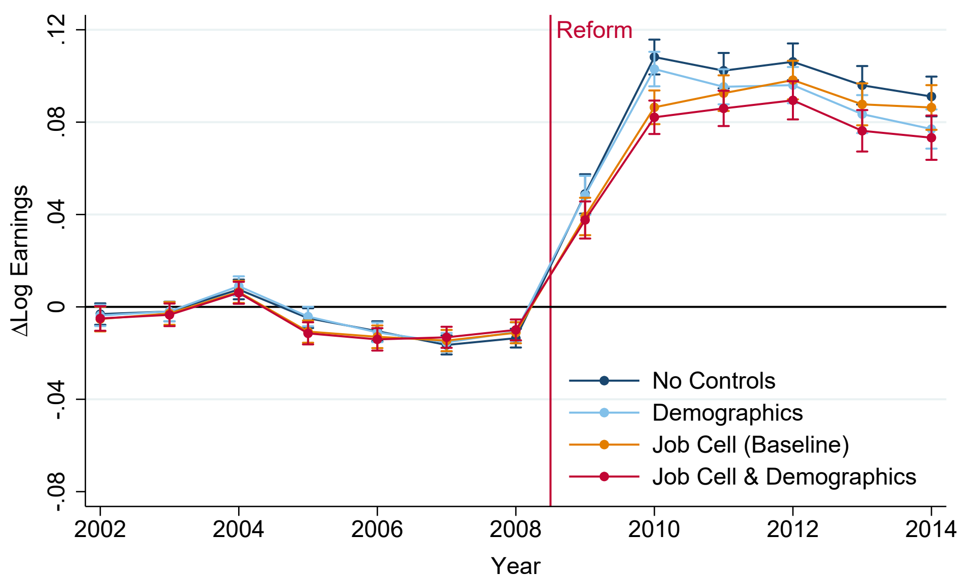
Figure A.7: Impact of Tax Reform on Earnings over Time

Robustness of Dynamic Approach to Controls

A: Intention-to-Treat (ITT)

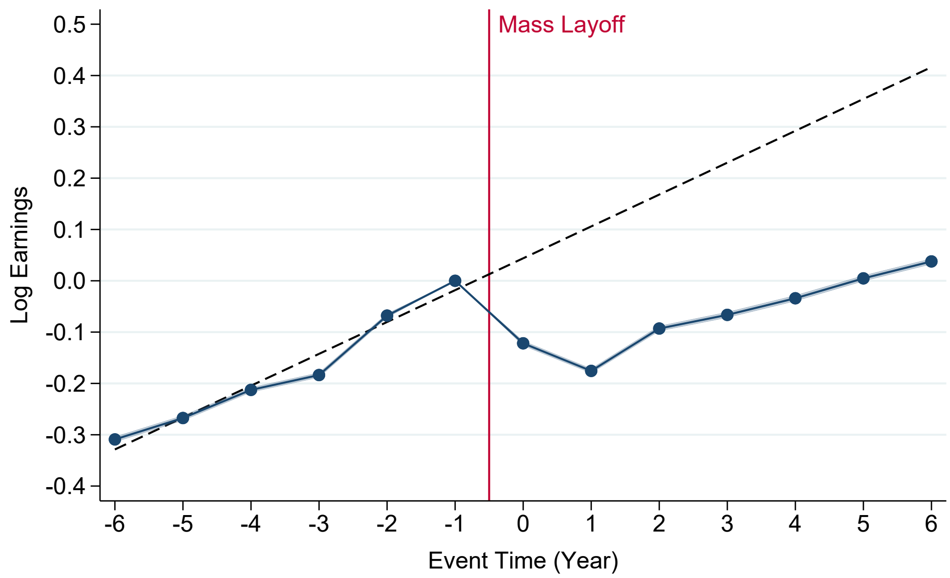


B: Treatment-on-the-Treated (TOT)



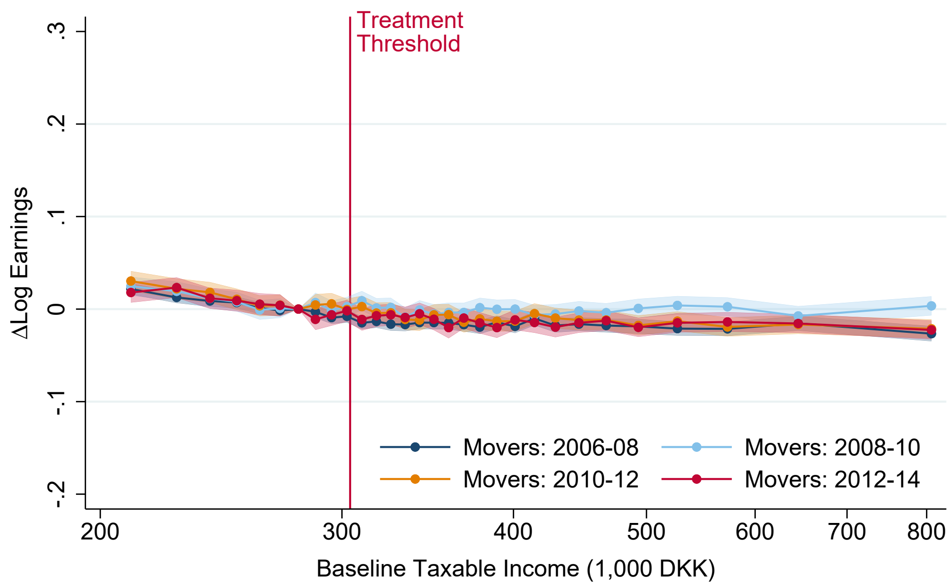
Notes: This figure investigates the robustness of the results in Figure 7. It plots the time profile of earnings effects (ITT and TOT, respectively) among job switchers under different sets of controls. Four different specifications are considered: no controls (dark blue), demographic controls (light blue), fixed effects for initial job cell (orange), and finally job cell fixed effects and demographic controls combined (red). The demographic variables include dummies for age, gender, marital status, and number of children. The specification with only job cell fixed effects corresponds to our baseline results in Figure 7. The analysis shows that our dynamic approach is robust to the specification of controls, consistent with the finding in Figure A.6 that job switching is not selected on observables. The error bars depict 95% confidence intervals based on robust standard errors.

FIGURE A.8: IMPACT OF MASS LAYOFF ON EARNINGS



Notes: This figure presents an event study of the effect of mass layoffs on earnings. Mass layoffs are defined as layoffs in which firms with at least 20 employees reduce their workforce by at least 30% in a single year. The figure shows log earnings by event time (blue series) compared to a linear time trend estimated on pre-layoff data (dashed line). Mass layoffs lead to sizeable and persistent earnings losses. The shaded area depicts 95% confidence intervals based on standard errors clustered at the individual level.

FIGURE A.9: IMPACT OF TAX REFORM ON FIRM-LEVEL WAGE PREMIA FOR SWITCHERS



Notes: This figure investigates if the earnings responses to lower taxes among job switchers are mediated by firm-level wage premia. To this end, we first estimate an AKM model of log earnings on individual fixed effects, firm fixed effects, and time-varying controls (dummies for year, age, and tenure). Restricting attention to firm switchers, we then regress the change in firm effects on baseline income bin, omitting a bin below the treatment threshold. This gives difference-in-differences estimates of the effect of the tax reform on firm-specific earnings premia for firm switchers. The figure plots these estimates over different time intervals: 2006-08 (placebo), 2008-10, 2010-12, and 2012-14. In every time interval and at all income levels, the coefficients are close to zero and (mostly) statistically insignificant. Hence, the earnings responses of firm switchers are not driven by tax-induced sorting into firms with higher wage premia. The shaded areas show 95% confidence intervals based on robust standard errors.

B Theoretical Proofs

B.1 Proof of Proposition 1

We insert flow utility (1) into the objective (3), which gives the following maximization problem:

$$\max_{y_t} \left\{ \sum_{s=t}^{\infty} \delta^{s-t} \mathbb{E}\left[(1-\tau) z_s\right] - n_t v\left(y_t/n_t\right) \right\}.$$

The first-order condition with respect to y_t is given by

$$(1-\tau) \sum_{s=t}^{\infty} \delta^{s-t} \frac{d}{dy_t} \mathbb{E}\left[z_s\right] = v'\left(y_t/n_t\right). \tag{29}$$

Using equation (2) to substitute for $\mathbb{E}\left[z_s\right]$, we obtain

$$\lambda (1-\tau) \sum_{s=t}^{\infty} \delta^{s-t} (1-\lambda)^{s-t} = v'\left(y_t/n_t\right).$$

Given the parameterization $v(x) = \frac{\eta}{\eta+1} x^{\frac{\eta+1}{\eta}}$, this may be rewritten as

$$\lambda (1-\tau) \sum_{s=t}^{\infty} \left(\delta (1-\lambda)\right)^{s-t} = \left(y_t/n_t\right)^{\frac{1}{\eta}}.$$

Finally, by using the relationship $\sum_{s=t}^{\infty} x^{s-t} = \frac{1}{1-x}$, we obtain the result in equation (4).

B.2 Proof of Proposition 2, Part 2

The correlation coefficient between z_t and y_t equals

$$\text{corr}(z_t, y_t) = \frac{\text{cov}\left(z_t, y_t\right)}{\sigma_{z_t} \sigma_{y_t}}, \tag{30}$$

where the covariance is defined as $\text{cov}(z_t, y_t) \equiv \mathbb{E}\left[(z_t - \bar{z}_t)(y_t - \bar{y}_t)\right]$. Using equation (2), this covariance may be written as

$$\begin{aligned}
\text{cov}(z_t, y_t) &= \mathbb{E}\left[\left(\lambda\left(y_t - \bar{y}_t\right) + (1-\lambda)\left(z_{t-1} - \bar{z}_{t-1}\right)\right)\left(y_t - \bar{y}_t\right)\right] \\
&= \mathbb{E}\left[\lambda\left(y_t - \bar{y}_t\right)^2 + (1-\lambda)\left(z_{t-1} - \bar{z}_{t-1}\right)\left(y_t - \bar{y}_t\right)\right] \\
&= \lambda \text{var}(y_t) + (1-\lambda)\,\text{cov}(z_{t-1}, y_t) \\
&= \lambda \text{var}(y_t),
\end{aligned}$$

where we used that $\text{cov}(z_{t-1}, y_t) = 0$, because y_t depends only on the current realization of n_t, while z_{t-1} only depends on realizations of n_s for periods $s < t$.

To compute the correlation coefficient, we also use that $\sigma^2_{z_t} = \lambda\sigma^2_{y_t} + (1-\lambda)\,\sigma^2_{z_{t-1}}$ from the earnings specification (2). This implies that $\sigma^2_{z_t} = \lambda\sum_{s=0}^{\infty}(1-\lambda)^s\,\sigma^2_{y_{t-s}}$. From equation (4) and $n_t = g(t) + \mu$, it follows that $\sigma^2_{y_t}$ is time-invariant, i.e. $\sigma^2_{y_t} = \sigma^2_y$ for $\forall t$. Using this time-invariance along with the property $\sum_{s=0}^{\infty} x^s = \frac{1}{1-x}$, it follows that $\sigma^2_{z_t} = \sigma^2_y$ and, hence, $\sigma_{z_t} = \sigma_y$. By inserting this property and the above formula for the covariance into the definition in (30), we obtain

$$\text{corr}(z_t, y_t) = \frac{\text{cov}\,(z_t, y_t)}{\sigma_{z_t}\sigma_{y_t}} = \frac{\lambda\sigma^2_{y_t}}{\sigma^2_{y_t}} = \lambda.$$

B.3 Social Welfare = Steady State Welfare When the Social Discount Factor is 1

Consider a social planner who wants to minimize the present discounted value of the deadweight loss from taxation, $\sum_{t=0}^{\infty}\rho^t D_t$, where ρ is the social discount factor. This objective is not well-defined for $\rho = 1$ and, therefore, we redefine the planner's objective function as

$$\Psi \equiv (1-\rho)\sum_{t=0}^{\infty}\rho^t D_t. \tag{31}$$

Because this objective function is just a monotone transformation of the original objective, they will yield identical optimal solutions. By adding and subtracting the steady state value D^*, the objective may be rewritten as

$$\begin{aligned}\Psi &= D^* + (1-\rho)\sum_{t=0}^{\infty}\rho^t\,(D_t - D^*)\\ &= D^* + (1-\rho)\sum_{t=0}^{T-1}\rho^t\,(D_t - D^*) + (1-\rho)\sum_{t=T}^{\infty}\rho^t\,(D_t - D^*)\end{aligned} \tag{32}$$

Given D_t is converging gradually towards D^*, the last term can be bounded:

$$\left|(1-\rho)\sum_{t=T}^{\infty}\rho^t\,(D_t - D^*)\right| \leq |D_T - D^*|\,(1-\rho)\sum_{t=T}^{\infty}\rho^t = |D_T - D^*|\,\rho^T.$$

By substituting this into equation (32), we obtain

$$\left|(1-\rho)\sum_{t=0}^{\infty}\rho^t D_t - D^*\right| \leq (1-\rho)\sum_{t=0}^{T-1}\rho^t\,|D_t - D^*| + |D_T - D^*|\,\rho^T \quad \forall T.$$

This implies

$$\lim_{\rho\to1}\left|(1-\rho)\sum_{t=0}^{\infty}\rho^t D_t - D^*\right| \le \lim_{\rho\to1}(1-\rho)\sum_{t=0}^{T-1}\rho^t\left(D_t-D^*\right)+\lim_{\rho\to1}\left|D_T-D^*\right|\rho^T \quad \forall T$$

$$\Leftrightarrow \quad \lim_{\rho\to1}\left|(1-\rho)\sum_{t=0}^{\infty}\rho^t D_t - D^*\right| \le \left|D_T-D^*\right| \quad \forall T.$$

Because D_T converges to D^* as T increases, it follows that

$$\lim_{\rho\to1}(1-\rho)\sum_{t=0}^{\infty}\rho^t D_t = D^*.$$

Therefore, at a social discount factor of $\rho = 1$, the welfare objective in equation (31) is equivalent to steady state welfare D^*. In this case, welfare analysis and policy design depend only on steady state elasticities, not the contemporaneous elasticities typically estimated.

B.4 Generalization of Proposition 3

When deriving equation (7), we disregarded any systematic lifecycle trend in earnings, i.e., $g(t)$ was assumed to be constant. In the general case where we impose only the initial condition $\bar{y}_0 = z_{-1}$, we obtain from equation (5):

$$\varepsilon_t^z = \frac{\lambda\sum_{s=0}^t(1-\lambda)^s\,\bar{y}_{t-s}\frac{d\bar{y}_{t-s}/\bar{y}_{t-s}}{d(1-\tau)/(1-\tau)}}{\lambda\sum_{s=0}^t(1-\lambda)^s\,\bar{y}_{t-s}+\left(1-\lambda\sum_{s=0}^t(1-\lambda)^s\right)\bar{z}_{-1}}.$$

From equation (4), we have $\frac{d\bar{y}_{t-s}/\bar{y}_{t-s}}{d(1-\tau)/(1-\tau)} = \eta$. Hence,

$$\varepsilon_t^z = \alpha_t\eta,$$

where

$$\alpha_t = \frac{\lambda\sum_{s=0}^t(1-\lambda)^s\,\bar{y}_{t-s}}{\lambda\sum_{s=0}^t(1-\lambda)^s\,\bar{y}_{t-s}+\left(1-\lambda\sum_{s=0}^t(1-\lambda)^s\right)z_{-1}}.$$

In this general expression, it remains the case that α_t increases over time from $\alpha_0 = \lambda$ to $\alpha_\infty = 1$.

B.5 Endogenous λ

If effort is observable or if workers can commit to an effort level, equilibrium earnings equal $y_t = (1-\tau)^\eta n_t$ as in a standard model. This maximizes worker-firm surplus (efficiency). We consider instead a setting where effort is unobservable without costly performance evaluations of workers. Evaluating a given worker costs q and reveals true effort y_t in the current period. Evaluations are carried out randomly with frequency λ. Considering a steady state with constant productivity n and effort y (to simplify exposition), we solve for the constrained-efficient solution of (y, λ) that maximizes worker-firm surplus.[41] The per-period surplus is given by

$$S = (1-\tau)\left[y - q\lambda\right] - nv\left(y/n\right),$$

where the term in square brackets is the net output/income generated. Note that, in this specification, we assume that evaluation costs $q\lambda$ are tax deductible. This will be the case if, for example, the costs of performance evaluations reflect labor costs.

The solution to y is still given by (4). The first-order condition for λ equals

$$\frac{dS}{d\lambda} = \left[1 - \tau - v'\left(y/n\right)\right]\frac{dy}{d\lambda} - q\left(1-\tau\right) = 0.$$

With costless verification ($q = 0$), we have $v'\left(y/n\right) = 1 - \tau$. Given the parameterization $v\left(x\right) = \frac{\eta}{\eta+1}x^{\frac{\eta+1}{\eta}}$ used previously, this implies $y = (1-\tau)^\eta n$ and is implemented by setting $\lambda = 1$ according to equation (4). With costly verification ($q > 0$), the incomplete information creates a wedge between the marginal benefit of effort $1-\tau$ and the marginal cost of effort $v'\left(y/n\right)$.

By inserting the marginal disutility of effort and using equation (4), we may rewrite the optimality condition as

$$\frac{dS}{d\lambda} = (1-\tau)\frac{(1-\lambda)(1-\delta)}{1-(1-\lambda)\delta}\cdot\frac{dy}{d\lambda} - q\left(1-\tau\right) = 0.$$

By differentiating equation (4) and rearranging terms, we obtain

$$\frac{dy}{d\lambda} = \frac{\eta\left(1-\delta\right)}{\lambda\left(1-\left(1-\lambda\right)\delta\right)}y,$$

[41] The solution can be decentralized in a competitive economy where workers receive compensation $(1-\tau)(y-f)$ where f equals $q\lambda$, which corresponds to firm spending on worker evaluations. In this situation, firm profits are zero in equilibrium.

which may be inserted into $dS/\partial\lambda = 0$ to arrive at the following equilibrium condition for λ:

$$\frac{\lambda}{1-\lambda} = \frac{\eta\left(1-\delta\right)^{2}}{\gamma\left(1-\left(1-\lambda\right)\delta\right)^{2}}, \tag{33}$$

where $\gamma \equiv q/y$ denotes the evaluation cost in proportion to output. We may interpret γ as capturing the degree/cost of imperfect information, which determines where λ lies in the interval between perfect verification ($\lambda = 1$ which obtains when $\gamma = 0$) and no verification ($\lambda = 0$ which obtains when $\gamma = \infty$). In general, for a positive and finite value of γ, the evaluation frequency λ lies between 0 and 1, thereby giving rise to the dynamic return mechanisms characterized in this paper. As for comparative statics, equation (33) shows that λ is decreasing in the evaluation cost γ, increasing in the effort elasticity η, decreasing in the discount factor δ, and independent of τ. The last result relies on the (natural) assumption that evaluation costs are tax deductible.

B.6 Derivation of Equations (15)-(16)

The expected profits of hiring a worker at time t on a fixed-wage contract $\hat{z}_t$ equals

$$\mathbb{E}\left[\pi\right] = \sum_{s=t}^{\infty} \delta^{s-t}\mathbb{E}\left[y_s - \hat{z}_t\right]\left(1-\lambda\right)^{s-t},$$

where $\left(1-\lambda\right)^{s-t}$ is the probability of retaining the worker in the same contract until time s. We assume free entry/exit of firms and that firms can pool risk. This implies that expected profits are zero in equilibrium. From the previous equation, we can solve for the competitive wage level as a function of expected worker output:

$$\hat{z}_t = \frac{\sum_{s=t}^{\infty} \delta^{s-t}\mathbb{E}\left[y_s\right]\left(1-\lambda\right)^{s-t}}{\sum_{s=t}^{\infty} \delta^{s-t}\left(1-\lambda\right)^{s-t}}. \tag{34}$$

The firm only observes worker output y_s at time $s = t$. Therefore, expected future output $\mathbb{E}\left[y_s\right]$ must align with the expected optimal choice of the worker. The first-order condition with respect to y_t is still given by equation (29). Using equation (14) to substitute for $\mathbb{E}\left[z_s\right]$ in equation (29) and noting that the wage will be fixed until the next job event occurs, we obtain

$$\lambda\left(1-\tau\right)\frac{d\mathbb{E}\left[\hat{z}_t\right]}{dy_t}\sum_{s=t}^{\infty}\delta^{s-t}\left(1-\lambda\right)^{s-t} = v'\left(y_t/n_t\right). \tag{35}$$

Since the productivity of the worker evolves according to $n_s = n_t (1+g)^{s-t}$, we guess that the solution is characterized by $y_s = y_t (1+g)^{s-t}$. We use this property to find a solution for $(y_t, \hat{z}_t)$ and then verify that the property is in fact satisfied for this solution. Using the property, we may write the wage equation (34) as

$$\hat{z}_t = y_t \frac{\sum_{s=t}^{\infty} \delta^{s-t} (1+g)^{s-t} (1-\lambda)^{s-t}}{\sum_{s=t}^{\infty} \delta^{s-t} (1-\lambda)^{s-t}} = y_t \frac{1-\delta(1-\lambda)}{1-\delta(1+g)(1-\lambda)}.$$

This is equation (16). From this equation, we also get

$$\frac{d\hat{z}_t}{dy_t} = \frac{1-\delta(1-\lambda)}{1-\delta(1+g)(1-\lambda)}.$$

By inserting this expression into the first-order condition (35), we obtain

$$\lambda(1-\tau)\frac{1-\delta(1-\lambda)}{1-\delta(1+g)(1-\lambda)}\sum_{s=t}^{\infty}\delta^{s-t}(1-\lambda)^{s-t} = v'(y_t/n_t),$$

which gives

$$\frac{\lambda}{1-\delta(1-\lambda)(1+g)}(1-\tau) = v'(y_t/n_t).$$

Given the parameterization $v(x) = \frac{\eta}{\eta+1} x^{\frac{\eta+1}{\eta}}$, this may be rewritten as

$$y_t = \left(\frac{\lambda}{1-\delta(1-\lambda)(1+g)} \cdot (1-\tau)\right)^{\eta} n_t.$$

This is equation (15). Finally, note that the solution for $(y_t, \hat{z}_t)$ characterized above satisfies the property $y_s = y_t (1+g)^{s-t}$ on which the derivations relied.

B.7 Derivation of Equation (18)

The first-order condition (29) still applies. Using equation (17) to substitute for $\mathbb{E}[z_s]$, we obtain

$$(1-\tau)[\lambda + (1-\lambda)\theta]\sum_{s=t}^{\infty}\delta^{s-t}(1-\lambda)^{s-t}(1-\theta)^{s-t} = v'(y_t/n_t)$$

Using the parameterization $v(x) = \frac{\eta}{\eta+1} x^{\frac{\eta+1}{\eta}}$ and the relationship $\sum_{s=t}^{\infty} x^{s-t} = \frac{1}{1-x}$, we obtain (18).

B.8 Generalized Earnings Dynamics for Movers and Stayers

The first-order condition (29) still applies. Using equation (20) to substitute for $\mathbb{E}[z_s]$, we obtain

$$(1-\tau)\left[\lambda\theta^m + (1-\lambda)\,\theta^s\right]\sum_{s=t}^{\infty}\delta^{s-t}\left[\lambda\,(1-\theta^m) + (1-\lambda)\,(1-\theta^s)\right]^{s-t} = v'\left(y_t/n_t\right)$$

Using the parameterization $v(x) = \frac{\eta}{\eta+1}x^{\frac{\eta+1}{\eta}}$ and the relationship $\sum_{s=t}^{\infty} x^{s-t} = \frac{1}{1-x}$, we obtain

$$y_t = \left(\frac{\lambda\theta^m + (1-\lambda)\,\theta^s}{1-\delta\left[\lambda\,(1-\theta^m) + (1-\lambda)\,(1-\theta^s)\right]}\cdot(1-\tau)\right)^{\eta} n_t,$$

showing that the elasticity of effort with respect to the net-of-tax rate is also η in this model version.

C The Role of Firm-Specific Wage Premia for Earnings Elasticities

Our approach to estimating earnings elasticities from job switchers uses variation from both firm and occupation transitions. As shown in Figure A.5, the earnings responses to lower taxes are similar for firm and occupation switchers. In this section, we focus on firm switchers and ask if their earnings responses are mediated by firm-level wage effects as studied in the literature on AKM models (Abowd, Kramarz, and Margolis 1999). That is, while our quasi-experimental estimates should be interpreted as worker responses (as they are based on tax variation across workers, not firms), they may be mediated by job switchers sorting into higher-wage firms following the tax reform. This would be a different mechanism than the one modeled in section 2, albeit consistent with our general emphasis on the importance of job switching for earnings responses.

To investigate the role of firm-level effects, we estimate a standard AKM model of the form

$$\log z_{it} = \alpha_i + \psi_{J(i,t)} + \boldsymbol{X_{it}\beta} + \nu_{it}, \tag{36}$$

where α_i is an individual fixed effect, $\psi_{J(i,t)}$ is a firm fixed effect, and $\boldsymbol{X_{it}}$ is a vector of time-varying controls. The controls include year dummies, age dummies, and dummies for tenure in the individual's current firm. We estimate the model using pre-reform data (2002-2005), restricting the sample to firms with at least 10 employees. We merge the estimated firm coefficients $\hat{\psi}_{J(i,t)}$ onto our tax reform sample, and regress the change in firm effects for job switchers on dummies for baseline income bin. This gives difference-in-differences estimates of the effect of the tax reform on firm-specific earnings premia for firm switchers by income bin. If the coefficients are positive in treated income bins, it implies that lower taxes induce switchers to sort into more remunerative firms, perhaps trading off non-wage amenities for higher wages.

The results are presented in Figure A.9. It plots the changes in firm-specific earnings premia by income bin in different time intervals: 2006-08 (placebo), 2008-10, 2010-12, and 2012-14. In every time interval and at all income levels, the coefficients are close to zero and statistically insignificant. In other words, the earnings responses for firm switchers are not driven by tax-induced sorting across firms with different wage premia. This is consistent with our theoretical model in which earnings responses reflect dynamic returns to individual effort, realized at the point of switching.