

# Micro vs Macro Labor Supply Elasticities: The Role of Dynamic Returns to Effort\*

Henrik Kleven<sup>†</sup>

Claus Kreiner<sup>‡</sup>

Kristian Larsen<sup>§</sup>

Jakob Søgaaard<sup>¶</sup>

August 2023

## Abstract

A key contention in economics is the discrepancy between micro and macro elasticities of labor supply with respect to marginal tax rates. We revisit this question, focusing on the role of dynamic returns to effort among top earners. 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. Short-run micro elasticities are attenuated relative to the true long-run macro elasticity. We proceed by providing two main empirical analyses using rich administrative data from Denmark. The first part presents descriptive evidence on earnings and hours-worked patterns over the lifecycle that confirm the predictions of the theoretical model. The second part presents quasi-experimental evidence on earnings responses to taxes using discrete job switches. The empirical strategy is informed by the theoretical model, according to which job switches can be used to (partially) identify the macro elasticity of labor supply. The evidence shows that, at the top of the distribution, macro elasticities are much larger than micro elasticities due to dynamic compensation effects.

---

\*We thank Pat Kline, Wojciech Kopczuk, Ilyana Kuziemko, Camille Landais, Felipe Lobel, Benjamin Lockwood, Emi Nakamura, Richard Rogerson, Benjamin Schoefer, Johannes Spinnewijn, Damián Vergara, Danny Yagan, and Owen Zidar for helpful comments and discussions. The activities of CEBI are financed by the Danish National Research Foundation, grant DNRF134.

<sup>†</sup>Princeton University, NBER, and CEBI. Email: kleven@princeton.edu.

<sup>‡</sup>University of Copenhagen and CEBI. Email: ctk@econ.ku.dk

<sup>§</sup>University of Copenhagen and CEBI. Email: kuol@econ.ku.dk

<sup>¶</sup>University of Copenhagen and CEBI. Email: 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 microeconomic evidence points to small elasticities, those skeptical of big government contend that the true long-run effect of taxes is large as they reduce the dynamism of labor markets (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).<sup>1</sup> This range is too large for economists to provide useful guidance on policy design.

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-term 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 potential importance of dynamic returns to effort. Consider an example close to home:

---

<sup>1</sup>These numbers are based on the Laffer rate formula  $\tau = 1 / (1 + \varepsilon a)$ , where  $\varepsilon$  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](#)). For example, the cross-country calibration study by [Prescott \(2004\)](#) argues for a (Hicksean) hours-worked elasticity much larger than 1.

top academics. An 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 implicitly rely on models where outcomes respond almost immediately to incentives such as in the static and frictionless model of hourly-paid workers.

Our agenda is complementary 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 estimation of models with human capital effects are consistent with large long-run elasticities (see e.g., [Keane 2011](#); [Keane and Rogerson 2015](#)).<sup>2</sup> 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. Our idea is related to a tradition in labor economics showing that job-to-job mobility is important for earnings growth either through gains in the job-match component of wages ([Topel and Ward 1992](#)) 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](#)). We develop a new model that highlights the role of dynamic returns realized at the point of switching and investigates 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 with a switching probability of one, in which case the elasticity of true effort  $\eta$  (which governs the long-run macro elasticity) is identical to the elasticity of realized earnings  $\varepsilon$  (the short-run micro

---

<sup>2</sup>[Best and Kleven \(2013\)](#) develop a theory of optimal taxation in a setting where effort affects future wages through changes in human capital accumulation.

elasticity). Outside this limit case, the macro elasticity is larger than the micro elasticity. Allowing for heterogeneity in both structural elasticities  $\eta$  and switching probabilities  $\lambda$ , we characterize the conditions under which the macro elasticity can be point identified or partially identified using responses among short-term switchers. The macro elasticity can be point identified when  $\eta$  and  $\lambda$  are orthogonal and partially identified when  $\eta$  and  $\lambda$  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 four facts of the data. First, earnings are strongly related to past hours worked, conditional on current hours worked. Considering workers at advanced career stages, past hours is a stronger predictor of earnings than current hours. Second, contemporaneous changes in hours worked and earnings are virtually unrelated at the top of the distribution, but not at the bottom. That is, while the hours-earnings relationship at the bottom of the distribution (such as for cleaners, cashiers, and hotel porters) is consistent with the textbook model of hourly-paid workers, the relationship higher up in the distribution (such as for engineers, lawyers, and managers) is consistent with a model of salaried workers for whom earnings and effort feature little correlation period by period. Third, 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 about 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. These empirical facts are consistent with our theoretical model.

Informed by our model and descriptive evidence, we provide a quasi-experimental study of earnings elasticities using firm $\times$ occupation switchers. The analysis is based on a recent tax reform experiment in Denmark: a reduction in the marginal tax rate above an income threshold located around the 70th percentile of the distribution. The tax rate reduction was large, about 10pp.<sup>3</sup>

---

<sup>3</sup>Given the tax reform was permanent, our estimates are most naturally interpreted as capturing Hicksian elasticities, not Frisch elasticities. A conceptual challenge to all tax reform studies, however, is that “permanent” tax changes are never truly permanent and behavioral responses may be influenced by (unobserved) expectations about future tax reforms. In this sense, it is conceivable that tax reform studies capture a mix of Hicksian and Frisch elasticities.

As in most of the modern literature on such reforms, behavioral responses are estimated using a difference-in-differences approach comparing treated and untreated workers from before to after the reform, supported by a set of transparent graphs. Considering the full population of treated workers, we find clear and precisely estimated earnings responses to taxes. The magnitude of these responses is modest, however, with an elasticity of earnings with respect to the net-of-tax rate of about 0.1. We then split the data into job movers and job stayers, showing that the small average response masks striking heterogeneity: job movers feature large responses — an elasticity of 0.4-0.5 — while job stayers feature precisely estimated zero responses. Because a minority of workers switch in any given year, the large switcher elasticity is consistent with a small average elasticity. We also consider the effect of the tax reform on the probability of switching, finding no effect on this margin. This is consistent with our theoretical model in which the switching probability, even when endogenously set by firms, does not respond to income taxes.

The central thesis of our paper is that earnings responses among short-run switchers can be used to uncover the long-run macro elasticity. As mentioned, point identification requires orthogonality between structural elasticities  $\eta$  and switching probabilities  $\lambda$ , the implication of which is that the observed elasticity  $\varepsilon$  is constant in the timing of switching. Estimating impacts by the timing of switching jobs, we find that early and late switchers feature very similar elasticities, consistent with point identification in our setting. We provide two additional pieces of evidence that speak to this point. First, we investigate if switcher characteristics respond to the reform — i.e., if any differences between switchers above and below the treatment threshold change from before to after the reform. Looking at a wide range of switcher characteristics, we find precisely estimated zero effects on all of them. As a result, controlling for switcher characteristics in the empirical specification hardly affects the estimates. Second, 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.<sup>4</sup> Taken together, this set of findings shows that our estimates are unlikely to be biased by selection in the decision to switch between firms or occupations.

Our paper contributes to several literatures in public finance, labor, and macroeconomics. First of all, we contribute to an enormous body of work estimating labor supply elasticities with respect

---

<sup>4</sup>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 entirely consistent with a negative reduced-form effect of a mass layoff itself (e.g., Jacobson, LaLonde, and Sullivan 1993).

to tax incentives, as reviewed by [Blundell and MaCurdy \(1999\)](#) and [Saez, Slemrod, and Gertz \(2012\)](#). Summarizing the microeconomic evidence, [Saez, Slemrod, and Gertz \(2012\)](#) argue that “the profession has settled on a value for this elasticity close to zero.”<sup>5</sup> Consistent with this view, we estimate a small earnings elasticity when taking a conventional quasi-experimental approach to studying tax cuts to top earners in Denmark.<sup>6</sup> We argue that such micro elasticities are uninformative of longer-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.<sup>7</sup> 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.

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

---

<sup>5</sup>To be clear, 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 feature much greater variation across studies and less of a consensus (see [Kleven 2023](#)).

<sup>6</sup>Studying the same Danish tax reform, [Kreiner, Leth-Petersen, and Skov \(2016\)](#), [Jakobsen and Sogaard \(2022\)](#), [Labanca and Pozzoli \(2022\)](#), and [Sigaard \(2022\)](#) also estimate small micro elasticities.

<sup>7</sup>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).

considerably smaller than those implied by a number of macro studies.<sup>8</sup>

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; 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. Given our empirical approach uses firm switchers, it is natural to ask if our earnings elasticities are mediated by firm-specific wage premia as estimated in the literature on AKM models (Abowd, Kramarz, and Margolis 1999). In other words, while our quasi-experimental estimates should be interpreted as worker responses (as they are based on tax variation across workers, not firms), these responses may be driven by workers moving into higher-wage firms. We estimate an AKM

---

<sup>8</sup>Our agenda is also complementary to a paper by Scheuer and Werning (2017) on the optimal taxation of superstars. 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.



model to investigate this point, showing that the earnings responses are not driven by such effects of taxes on worker sorting across firms. This is consistent with our theoretical model of dynamic returns to individual effort, realized at the point of switching.

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.

## 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

$$u_t = (1 - \tau) z_t - n_t v(y_t / n_t), \quad (1)$$

where  $\tau$  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  $\mu$  is an idiosyncratic, random shock.

The quasi-linear utility specification in (1) is commonly used in the literature on income taxation (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  $z_t = y_t = (1 - \tau)^\eta n_t$ , where  $\eta$  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  $\lambda$  in any given period. These job events align realized earnings with latent earnings (effort). Hence, we have



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

The basic idea is that worker effort is unobservable to employers absent costly performance evaluations, 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  $\lambda$  is exogenous, but we later develop a generalization in which the switching probability is endogenously set by firms facing effort verification costs. In either case, the value of  $\lambda$  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  $\lambda$  is small, the return to effort materializes far into the future in expectation.<sup>9</sup>

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  $\delta$ , the optimization problem can be written as

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

subject to the earnings dynamics in (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*

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

where  $\eta$  is the Hicksian 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  $\frac{\lambda}{1 - (1 - \lambda)\delta}$ , which 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 the level

---

<sup>9</sup>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 that the earnings of a given worker are adjusted, independently of the time since the last adjustment. Importantly, our objective — understanding how earnings and effort respond to taxes — is fundamentally different from the focus in macroeconomics on nominal price rigidity and the impact of monetary policy.

of effort. On the one hand, a lower  $\lambda$  implies that increased effort today is less likely to yield increased earnings today (short-run effect captured by the numerator of the adjustment factor). On the other hand, a lower  $\lambda$  implies that any increase in earnings is expected to last a longer period of time as the expected duration of a given job spell equals  $1/\lambda$  (long-run effect captured by 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. In fact, none of the key results derived below depend on the size of the discount factor.

Using equation (2), we can write average earnings at time  $t$  as a function of the average levels of effort from time 0 and the initial level of average earnings at time 0. We have

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

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

To summarize, the model has the following predictions on the variation in effort and realized earnings over time:

**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  $\lambda$ .*
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 dynamic compensation model has important implications for earnings responses to tax policy 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,  $\bar{z}_t = \bar{z}$ .<sup>10</sup> 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  $t$  with respect to the net-of-tax rate as  $\varepsilon_t^z \equiv \frac{d\bar{z}_t/\bar{z}_t}{d(1-\tau)/(1-\tau)}$ , the ratio of behavioral and mechanical effects can be written as

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

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  $\varepsilon_t^z$  is a sufficient statistic, conditional on  $\tau$ . The issue is that, with dynamic compensation,  $\varepsilon_t^z$  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.<sup>11</sup>

Computing the long-run welfare effect,  $dD_\infty$ , requires information about the long-run earnings elasticity,  $\varepsilon_\infty^z$ . 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  $\tau$  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.*

<sup>10</sup>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 an appendix, as discussed below.

<sup>11</sup>See Appendix B.3 for a proof.

Specifically,

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

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  $\lambda$ . (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  $\varepsilon_t^y \equiv \frac{d\bar{y}_t/\bar{y}_t}{d(1-\tau)/(1-\tau)}$ . 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  $\varepsilon_t^z$  underestimate the welfare-relevant, long-run elasticity  $\eta$ . The bias is increasing in the degree of dynamic compensation (inversely related to  $\lambda$ ). 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 derivations above disregarded lifecycle trends in earnings ( $g(t)$  was assumed to be constant), the results are generalized to allow for such lifecycle dynamics in Appendix B.4. There we show that the formula for  $\alpha_t$  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  $\mu$ , but all other parameters of the model were assumed to be homogeneous across workers. Realistically, there will also be heterogeneity in effort elasticities  $\eta$  and switching probabilities  $\lambda$ . 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.

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

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.

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

where  $f(\eta, \lambda | J_t = 1)$  denotes the density of  $\eta, \lambda$  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  $\tau$  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*

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

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

*Proof.* From Bayes' Rule, we have  $f(\eta, \lambda | J_t = 1) = \frac{P(J_t=1|\lambda)f(\eta,\lambda)}{P(J_t=1)}$ , where  $P(J_t = 1 | \lambda) = \lambda(1 - \lambda)^t$  and  $P(J_t = 1) = \mathbb{E}[\lambda(1 - \lambda)^t]$  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

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

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

Under orthogonality between effort elasticities  $\eta$  and switching probabilities  $\lambda$ , the macro elasticity can be point identified. Importantly, this case is associated with an observable signature in the data, namely that the switcher elasticity  $\mathbb{E}[\varepsilon_t^z | J_t = 1]$  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  $\eta$  and  $\lambda$  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  $t$ ,  $P(J_t = 1 | \lambda) = \lambda(1 - \lambda)^t$ , is increasing in  $\lambda$  for  $t < \frac{1-\lambda}{\lambda}$  and decreasing in  $\lambda$  for  $t > \frac{1-\lambda}{\lambda}$ . Therefore, if*

$\text{cov}(\eta, \lambda) > 0$ , we have that  $\text{cov}(\eta, \lambda(1 - \lambda)^t)$  is positive at  $t = 0$  and turns negative at a sufficiently large  $t$ . From equation (10), this implies that short-run switcher elasticities,  $\mathbb{E}[\varepsilon_t^z | J_t = 1]$  for small  $t$ , provide upward-biased estimates of the average long-run elasticity  $\mathbb{E}[\eta]$ . In this case, with estimates of short-run switcher elasticities for periods  $t = 0, \dots, T$ , a lower bound is given by

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

where  $\Lambda_t$  denotes the share of workers making their first post-reform job switch at time  $t$ . Conversely, if  $\text{cov}(\eta, \lambda) < 0$ , then  $\text{cov}(\eta, \lambda(1 - \lambda)^t)$  is negative at  $t = 0$  and turns positive at a sufficiently large  $t$ . From equation (10), this implies that  $\mathbb{E}[\varepsilon_0^z | J_0 = 1] < \mathbb{E}[\eta]$  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 Endogenizing $\lambda$

Appendix B.5 develops an extension of our model with an endogenous switching probability  $\lambda$ . 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  $\lambda$ . 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

$$S = (1 - \tau) [y - q\lambda] - nv(y/n), \quad (13)$$

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  $\lambda$ ).** *With a positive and finite evaluation cost  $q$ , the equilibrium evaluation frequency  $\lambda \in (0, 1)$ . The limit case of perfect verification ( $\lambda = 1$  as in the standard model) is obtained for  $q = 0$ , while the limit case of no verification ( $\lambda = 0$ ) is obtained for  $q = \infty$ . Outside these limit cases, we have that  $\lambda$  is decreasing in the evaluation cost  $q$ , increasing in the effort elasticity  $\eta$ , decreasing in the discount factor  $\delta$ , and independent of  $\tau$ .*

*Proof.* See Appendix B.5. ■

This proposition microfound the dynamic compensation model and implies that all of our previous results generalize. Two specific 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, heterogeneity in  $\lambda$  would be driven by underlying heterogeneity in evaluation costs  $q$ . The identification results of the previous section depend on the correlation between  $\eta$  and  $\lambda$ , which would be driven by the joint density of  $(\eta, q)$  and the fact that  $\lambda$  is directly increasing  $\eta$  by Proposition 5. Due to the direct effect,  $\eta$  and  $\lambda$  would tend to be positively correlated, in which case short-run switcher elasticities partially identify the long-run elasticity, with the lower bound given in equation (12).

### 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. 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, all 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. The table reports average earnings across all workers in each occupation cell and across workers in the top 1% largest firms (in terms of number of employees).

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.<sup>12</sup> Still, given hours worked is a component of effort, we provide descriptive evidence 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  $h_m$  is divided into four bins.<sup>13</sup> 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.<sup>14</sup> 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

### 4.1 Four Descriptive Facts

To validate our theoretical model of dynamic compensation, this section presents descriptive facts that speak to the predictions of the model. The analysis leverages the granularity and statistical power of the Danish data to provide particularly clear evidence. As we shall see, the evidence is consistent with the dynamic compensation model *and* inconsistent with the standard labor supply model. The findings lend support to the quasi-experimental approach developed later.

---

<sup>12</sup>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 reason is that unobserved effort choices influence the quality-adjusted hours relevant for earnings progression.

<sup>13</sup>This gives a total of 37 hours bins ( $= 4 \times 12 - 12 + 1$ ) over a year.

<sup>14</sup>Even so, Figure A.1 in the appendix shows that the administrative pension measure of hours worked is very good. The figure compares the relationship between earnings and hours worked in the administrative data (Panel A) to the relationship in labor force survey data (Panel B). The survey data contain information on self-reported hours worked (actual, uncapped hours). Reassuringly, the earnings-hours relationship is similar in the two data sources. But 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.

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

Any model of 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 supported 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.<sup>15</sup> Each panel shows the non-parametric relationship between the average earnings rank at age 50 and hours 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 worked may reflect that variation in past hours captures variation in current hours. In addition, variation in past hours may be correlated with variation in productivity 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 both current hours and demographic variables that proxy for earnings capacity (red dots).<sup>16</sup> The controls dampen the correlation between earnings and past hours as one would expect, but the relationship remains strong. The expected earnings rank at age 50, conditional on current hours and demographics, increases from the 25th percentile to the 65th percentile as annual hours worked over the preceding 10 years increases from zero to 2,000 hours. The relationship becomes 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.

These findings suggest that the return to effort (hours) has a strong dynamic component. The evidence contradicts standard labor supply models in which current hours of work is a sufficient statistic for earnings.

---

<sup>15</sup>In this analysis, because we consider a long panel, working hours are measured using employer pension contributions (ATP) available across all years of the data, as described in Section 3.

<sup>16</sup>The demographic controls include dummies for education level (8 categories), gender (binary), children (binary), and marital status (7 categories).

#### 4.1.2 Fact 2: Hours and Earnings Changes are Contemporaneously Unrelated at the Top

Another prediction of the dynamic compensation model is that the contemporaneous correlation between hours and earnings changes is small.<sup>17</sup> This stands in contrast to standard models of hourly-paid workers in which the two are perfectly correlated. To shed light on the two models, Figure 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 shown by blue dots, while examples of specific occupations are shown by red triangles and diamonds. To capture hours variation even among top earners, these graphs use monthly information on uncapped hours available since 2008.<sup>18</sup> Importantly, the theoretical prediction is about the intensive margin, not the extensive margin. To avoid effects from the extensive margin, we consider average monthly outcomes over the year calculated across months with positive hours and earnings. Our results therefore capture correlations at the annual level, neutralizing any variation from the number of months worked.

As can be seen from the figure, there is a stark difference between workers at the bottom and at the top of the earnings distribution. At the bottom, the relationship between log hours and log earnings is very close to the 45-degree line, as predicted by standard models of hourly-paid workers. This is natural when considering typical occupations at the bottom: the examples provided in the figure — cleaners and waiters — are hourly-paid jobs. At the top of the distribution, on the other hand, the relationship between log hours and log earnings is much flatter. The slope remains positive (albeit small) in the top 20% and top 10%, while it is virtually zero in the top 1%. Again, this is natural when considering typical occupations at the top. All of the examples shown — pharmacists, engineers, lawyers, CEOs, etc. — are salaried jobs that have no immediate link between working hours and earnings. The contemporaneous correlation is not quite zero, consistent with our theoretical model in which the probability of switching jobs creates a within-period link between effort and earnings. The fact that the correlation decreases as we move further into the top tail may be driven by heterogeneity in job switching probabilities across different quantiles of earnings. As workers reach the very top, switching across occupations or firms becomes less frequent.

These findings suggest that, while the standard labor supply model may be a reasonable ap-

---

<sup>17</sup>Recall that, in our model, the contemporaneous correlation coefficient between hours and earnings is equal to the unconditional job switching probability  $\lambda \ll 1$ .

<sup>18</sup>See section 3 for details.

proximation for the bottom of the labor market, it is a poor description of the top of the labor market. For the top, the patterns are instead consistent with the dynamic compensation model.

#### 4.1.3 Fact 3: Lifecycle Profiles in Earnings are Driven by Discrete 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 borne out by the data. This is critical for understanding how to model labor supply and for developing an empirical strategy to estimate long-run responses to taxes.

Leveraging the granularity of the Danish data, we measure job switches as transitions between firm $\times$ occupation cells. The first set of results is presented in Figure 3. 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 $\times$ firm fixed effects (orange), and net of occupation $\times$ firm $\times$ 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: about 95% 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.2 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 12 at the top (11 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

$$z_{ijt} = \bar{z}_{ij} + (z_{ijt} - \bar{z}_{ij}), \quad (14)$$

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

$$\text{var}(z_{ijt} - \bar{z}_i) = \underbrace{\text{var}(\bar{z}_{ij} - \bar{z}_i)}_{\text{Between Job}} + \underbrace{\text{var}(z_{ijt} - \bar{z}_{ij})}_{\text{Within Job}}, \quad (15)$$

where we use that  $\text{cov}(\bar{z}_{ij} - \bar{z}_i, z_{ijt} - \bar{z}_{ij}) = 0$ . Hence, the variance of individual earnings over the lifecycle (relative to its mean) can be decomposed into between-job and within-job variances, with no covariance term.

Figure 4 presents the results of such an analysis. Considering the same panel of workers analyzed 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 about 95% 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.

#### 4.1.4 Fact 4: 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 a 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.<sup>19</sup> 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

$$Y_{iq} = \sum_j \alpha_j \cdot Event_{j=q} + \beta \cdot Treat_i + \sum_{j \neq -1} \gamma_j \cdot Event_{j=q} \cdot Treat_i + \phi_{q \in t} + \phi_a + \nu_{iq}, \quad (16)$$

where  $Event_{j=q}$  is a quarterly event time dummy,  $Treat_i$  is a promotion dummy,  $\phi_{q \in t}$  is a calendar year fixed effect, and  $\phi_a$  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  $\gamma_q$ . 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 5 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 5% 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

---

<sup>19</sup>In selecting the sample, individuals are required to stay in the same firm from two quarters before promotion to two quarters after promotion.

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 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

To obtain exogenous variation in tax rates, we use a major tax reform implemented in Denmark in 2009-10.<sup>20</sup> 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, raised the top bracket threshold, and reduced the marginal tax rate within the top bracket. 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 by the reform. The threshold that separates treatments and controls is located at around the 70th percentile of the income distribution. Figure 6 shows marginal tax rates over time for taxpayers above the treatment threshold (red series) and taxpayers below the treatment threshold (blue series). The reform-induced reduction in the top marginal tax rates was large, almost 10 percentage points. We note that, while the reform-induced tax variation 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, our objective is to demonstrate our approach to estimating long-run macro elasticities using a widely available source of tax variation.

The income concept that determines treatment assignment includes labor income, transfers, pensions, alimony, and certain capital income items. We divide this taxable income measure into a set of discrete bins  $b$  around the treatment threshold, omitting a bin  $b_0$  below the threshold. We

---

<sup>20</sup>We refer to [Jakobsen and Sogaard \(2022\)](#) for a detailed description of the Danish tax system and 2009 reform.



consider a difference-in-differences specification of the form

$$\Delta \log z_{it} = \beta_0 + \sum_{b \neq b_0} \beta_b \cdot \mathbf{1}[y_{i0} \in b] + \sum_{b \neq b_0} \gamma_b \cdot \mathbf{1}[\text{switch}_{it}] \cdot \mathbf{1}[y_{i0} \in b] + \nu_{it}, \quad (17)$$

where  $z_{it}$  denotes the labor income of individual  $i$  at time  $t$ ,  $y_{i0}$  denotes the taxable income of individual  $i$  prior to the reform, and  $\mathbf{1}[\cdot]$  denotes an indicator function. Based on the model developed in section 2, long-run responses to tax changes can be estimated from short-run responses among job switchers. We therefore interact treatment status (pre-reform income bin) with an indicator for switching job, defined as moving between firm  $\times$  occupation cells.<sup>21</sup> This allows for estimating difference-in-differences coefficients separately for stayers ( $\hat{\beta}_b$  in income bin  $b$ ) and movers ( $\hat{\beta}_b + \hat{\gamma}_b$  in income bin  $b$ ). By estimating these coefficients by income bin, we are able to investigate heterogeneity in behavioral responses (for movers vs stayers) across different income levels.

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  $y_{i0}$ . The most obvious confounder is mean reversion, as discussed extensively in the 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 Sogaard \(2022\)](#), we address this issue by running our regression in pre-reform and post-reform datasets separately. 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  $P$ , behavioral responses can be estimated based on a triple-differences approach:  $\hat{\beta}_b - \hat{\beta}_b^P$  for stayers and  $\hat{\beta}_b + \hat{\gamma}_b - (\hat{\beta}_b^P + \hat{\gamma}_b^P)$  for movers. Alternatively, we may take a quadruple-differences approach by considering the difference between the triple-differences for movers and stayers, i.e.  $\hat{\gamma}_b - \hat{\gamma}_b^P$ .

We convert the estimates of earnings responses into elasticities with respect to  $1 - \tau$ . For this purpose, we pool the income bins in equation (17) into treated and untreated ranges (above and below the threshold), re-estimating the equation with these broader treatment categories. In general, reduced-form estimates that use pre-reform behavior to assign treatment status will be attenuated

---

<sup>21</sup>Specifically,  $\mathbf{1}[\text{switch}_{it}]$  equals one for workers who make their first post-reform job switch at time  $t$ .

due to people moving across the treatment threshold after the reform. To alleviate this issue, the elasticity calculations are based on a donut-hole approach in which we drop observations with pre-reform incomes close to the threshold.<sup>22</sup> Denoting the pooled regression estimates by  $\hat{\beta}$ ,  $\hat{\beta}^P$ ,  $\hat{\gamma}$ , and  $\hat{\gamma}^P$ , the elasticity for job movers can be calculated as

$$\varepsilon|_{\text{switch}} = \frac{\hat{\beta} + \hat{\gamma} - (\hat{\beta}^P + \hat{\gamma}^P)}{\mathbb{E} [\Delta \log (1 - \tau_{it}) | T] - \mathbb{E} [\Delta \log (1 - \tau_{it}) | C]}, \quad (18)$$

when using the triple-differences specification described above. Alternatively, when using the quadruple-differences specification (triple-differences for movers relative to stayers), the numerator of the elasticity is set equal to  $\hat{\gamma} - \hat{\gamma}^P$ . The denominator equals the average change in  $\log (1 - \tau_{it})$  from before to after the reform for treatments ( $T$ ) relative to controls ( $C$ ). As shown in Figure 6, the tax change in the control group is close to zero.

As mentioned, estimating behavioral responses by comparing individuals in different pre-reform tax brackets is associated with attenuation bias. Some taxpayers assigned to the treatment (control) group have incomes below (above) the treatment threshold after the reform. Dropping observations close to the threshold alleviates the issue, but does not eliminate it. As a result, the reduced-form estimates described above should be interpreted as intention-to-treat (ITT) effects. Conceptually, we are more interested in treatment-on-the-treated (TOT) effects. To estimate TOT elasticities, we specify the pooled version of equation (17) in terms of current taxable income  $y_{it}$ , constructing instruments based on pre-reform taxable income  $y_{i0}$ . To ensure that pre-reform tax bracket is a strong predictor of post-reform tax bracket, we continue to drop observations close to the threshold. The model is estimated using 2SLS. The elasticity is still defined as in equation (18), only the numerator is based on TOT coefficients obtained from the 2SLS estimation.

As shown in section 2, for the switcher elasticity to point identify the long-run elasticity of interest, the timing of switching cannot be selected on the underlying structural elasticity. We provide several analyses to rule out such selection. First, we estimate if the switcher elasticity varies by the timing of switching jobs following the reform. We find that the elasticity is very stable over time, consistent with our theoretical prediction under orthogonality between switching probabilities and structural elasticities. Second, we investigate if switcher characteristics respond to the reform, using the same empirical design as we use for estimating earnings responses. That is, we ask if any observable differences between switchers above and below the treatment threshold

---

<sup>22</sup>Specifically, we drop observations located within 30,000 DKK of the treatment threshold, corresponding to about 10% of threshold income.

change from before to after the reform. Looking at a wide range of characteristics, we find precisely estimated zero effects on all of them. As a result, controlling for demographic variables in specification (17) hardly affects the estimates. Finally, we restrict the sample to plausibly exogenous switches, namely those triggered by mass layoffs. We show that mass-layoff switchers feature similar earnings responses as the full sample of switchers. All of these results are supported by transparent, non-parametric graphical evidence.

## 5.2 Impact of Tax Reform: Switchers vs Non-Switchers

### 5.2.1 Reduced-Form Effects

We start by investigating earnings responses to the 2009 tax reform in the full sample of treated workers. This analysis is based on a simplified version of specification (17) without the interaction between pre-reform income and job switching dummies. The results are presented in Figure 7. It plots the changes in log earnings between 2008-10 (actual experiment) and between 2006-08 (placebo experiment) by income bin. The threshold above which marginal tax rates were reduced is depicted by the red vertical line. The 2008-10 series shows that earnings growth above the treatment threshold is *smaller* than below the threshold, implying that a standard difference-in-differences approach would yield negative elasticities. As discussed above, this is likely to reflect downward bias from mean reversion: those with high pre-reform incomes tend to be selected on positive transitory shocks, reducing their earnings growth over time regardless of the tax cuts. Comparing the 2008-10 series to the 2006-08 series addresses this issue. The two series track each other very closely below the treatment threshold — consistent with mean reversion being stable over time — but diverge above the threshold. There is a clear and statistically significant earnings response in every bin above the threshold. As we show later, however, the behavioral response is modest in elasticity terms.

Having established the presence of behavioral responses in the full sample, we investigate heterogeneity by job switching status in Figure 8. As in the previous figure, we show earnings growth between 2008-10 and 2006-08 by income bin, but do so separately for switchers (Panel A) and non-switchers (Panel B). Switchers are those who move across firm $\times$ occupation cells between 2008-10 and 2006-08, respectively. The patterns are striking: the earnings responses among job movers are large — roughly four times as large as in the full sample — whereas the responses among job stayers are close to zero and statistically insignificant. These results are consistent with

our dynamic compensation model in which earnings responses to taxes are realized only at the point of switching jobs. It is also worth noting that, among job movers, the earnings responses are increasing in the level of income. This suggests that dynamic compensation effects increase in importance as we move further into the top tail of the distribution.

Do taxes impact the probability of switching? As shown in section 2.4, even when switching probabilities are endogenously determined based on the costs of verifying true effort, they do not respond to taxes as long as verification costs are tax deductible. Figure 9 investigates if this prediction is borne out by the data. The figure is constructed in the same way as the previous figures, but replaces the dependent variable with an indicator for switching jobs. The tax reform has no effect on the probability of switching: the relationship between the switching probability and taxable income is exactly the same before and after the reform. Workers below the treatment threshold are more likely to switch, but this is equally true between 2008-10 and between 2006-08. This finding, besides confirming one of the predictions of our model, alleviates concerns that job switching is selected. If the tax reform had increased the likelihood of switching, we would expect the marginal switchers — presumably those with the strongest incentive to change jobs when taxes are lower — to be selected on large elasticities. The absence of tax-induced job switching rules out such endogeneity. Switching could still be selected, but Figure 9 provides an important first step in ruling it out.

Does the earnings impact of taxes on job switchers vary by the timing of switching? We have seen that initial job movers feature large responses while initial job stayers feature no responses, but the initial stayers eventually become movers and will reveal their responses at that time.<sup>23</sup> The fundamental 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 the return is not realized until the time of moving. To shed light on this idea, we estimate the earnings responses of job movers at different points in time. To interpret this time profile, it is important to convert the intent-to-treat (ITT) effects into treatment-on-the-treated (TOT) effects. As discussed in the preceding section, ITT estimates based on comparing individuals in different pre-reform tax brackets are associated with attenuation bias due to people moving across the bracket threshold after the reform. Such attenuation becomes stronger over time, thus confounding the interpretation of the time profile in ITT effects. We estimate TOT effects based on a 2SLS specification in which the actual post-reform tax bracket is instrumented using the pre-reform tax bracket.

---

<sup>23</sup>As shown in Appendix Figure A.2, all workers switch jobs — and typically many times — over their careers.

The results are presented in Figure 10. The figure plots the average earnings impact on job switchers over increasingly long time intervals: 2008-10, 2008-11, 2008-12, 2008-13, and 2008-14. The black series depict ITT effects while the red series depict TOT effects. All estimates are based on the quadruple-differences specification described in section 5.1. The TOT series lie above the ITT series, reflecting the aforementioned attenuation bias in ITT estimates. Importantly, the TOT effects on job switchers are very stable over time. Hence, the earnings responses among workers who switch immediately after the reform and workers who switch later are about the same.

These findings are important for assessing causal identification in our empirical design. As shown in section 2, point identification of the long-run elasticity using short-run switchers requires orthogonality between structural elasticities and switching probabilities. This is associated with an observable pattern in the data, namely that the earnings response by switchers is constant in the timing of switching. Conversely, if structural elasticities and switching probabilities were correlated, the earnings response by switchers would be either declining (positive correlation) or increasing (negative correlation) as we consider switches farther removed from the time of the reform. The results presented here are consistent with point identification.

### 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. The table presents estimates based on the standard approach (top panel) and the dynamic approach (bottom panel). As described above, the standard approach is to estimate the average effect of taxes on all workers (applying our triple-differences specification) while the dynamic approach is to estimate the effect of taxes on job movers relative to job stayers (applying our quadruple-differences specification). The table shows both intention-to-treat (ITT) and treatment-on-the-treated (TOT) estimates, and it shows how these estimates vary over time. The estimates of primary interest — highlighted by boldface in the table — are the are the TOT elasticities based on the dynamic approach.

The table provides two main insights. First, the standard approach yields elasticities which are small and unstable over time. The short-run elasticity equals 0.1 — a conventional magnitude for micro elasticities — but the elasticity declines over time and eventually turns negative. This highlights the difficulties of trying to estimate longer-run elasticities simply by extending the event time window around tax reforms. The reason is that the threat from non-tax confounders becomes more serious over time. Specifically, the negative estimates in the top panel reflects bias in the

adjustment for mean reversion (based on a pre-reform placebo difference-in-differences) over long time windows. Second, the dynamic approach yields elasticities which are large and stable over time. The elasticities fall in a narrow band of 0.44-0.49. Viewed through the lens of our model, the constancy of switcher elasticities implies orthogonality between structural elasticities and switching probabilities, in which case the long-run elasticity is point identified. As an empirical matter, the reason why the dynamic approach is less vulnerable to bias from mean reversion as we extend the time window is the fact that we use a quadruple-differences specification in which we compare job movers and job stayers. Because stayers are unresponsive to the tax reform, they provide a within-period handle on mean reversion and other non-tax confounders.

Figure 11 investigates heterogeneity in elasticities by income level. It plots elasticities based on the standard approach (black series) and the dynamic switcher approach (red series) above different income cutoffs, depicted on the x-axis. All elasticities are based on TOT effects between 2008-10. The figure shows that the standard elasticity is small and virtually constant in income, whereas the dynamic elasticity is large and increasing in income. The dynamic elasticity increases from about 0.45 when considering all switchers above the treatment threshold (corresponding to the estimates of average elasticities in the table) to about 0.6 when considering switchers at the very top of the distribution.<sup>24</sup> The fact that the gap between dynamic and standard elasticities increases in income suggests that dynamic compensation effects become more important as we move further into the upper tail of the distribution.

### 5.3 Identification: Is Switching Selected?

The preceding analysis presented several findings that lend support to our identification strategy. This includes the findings that switching probabilities did not respond to the tax reform and that earnings responses did not decline for later switchers. Both of these results are consistent with the idea that short-term switcher elasticities can be used to uncover long-run elasticities. In this section, we present a number of additional identification checks, focusing on whether job switching is selected in a way that could bias our estimates. The first check investigates behavioral responses to a placebo reform. The results are presented in Figure 12, which plots earnings growth by income bin between 2002-04 and 2004-06, both before the 2009 tax reform experiment. Reassuringly, we find no earnings responses to this placebo experiment among either movers or stayers. The

---

<sup>24</sup>The highest income cutoff shown in the figure (700,000 DKK) corresponds approximately to the 98th percentile of the income distribution.

2002-04 and 2004-06 series track each other closely above and below the treatment threshold for both groups.

The next analysis investigates if job switching is selected on observables given our empirical design. Specifically, treating demographic characteristics as dependent variables in our quasi-experimental specification, we ask if any of them respond to the reform? The evidence presented in Figure 13 shows that the answer is a resounding no. The figure plots demographic variables by income bin for workers who switch jobs between 2008-10 and 2006-08, respectively. We consider six variables that matter for labor supply: age, fraction male, fraction married, number of children, fraction college educated, and fraction being manager. These variables are increasing in income — thus being higher among treatments than among controls — but the relationship is virtually identical before and after the reform. In other words, there is nothing observably different about treated relative to untreated switchers after the reform compared to before the reform. As a result, adding demographic controls to the regression equation (17) should not make any material difference to the estimates. This is confirmed in Appendix Figure A.3.

An alternative approach is to focus on job switches triggered by a plausibly exogenous event. A large literature on the effects of job displacement uses mass layoffs as a source of exogenous variation (e.g., [Jacobson, LaLonde, and Sullivan 1993](#)). Building on this literature, we implement our empirical approach in the sample of workers who switch firms following a mass layoff. The findings are presented in Figure 14. 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.<sup>25</sup> We find that the earnings responses to lower taxes are large even among mass-layoff movers. In fact, the responses are slightly 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.<sup>26</sup>

Taken together, the analyses presented in this section suggest that our estimates of switcher

---

<sup>25</sup>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 the 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.

<sup>26</sup>The positive earnings effect of lower taxes among mass-layoff switchers is consistent with a negative earnings effect of the mass layoff itself, as documented in the literature on the effects of job displacement. To confirm this, Figure A.4 in the appendix provides an event study of mass layoffs, showing that they generate sizeable earnings losses. The findings in Figure 14 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 such tax cuts.



elasticities are not biased, or at least not *upward* biased, by selection into job switching.

#### 5.4 Are Earnings Elasticities Mediated by Firm-Specific Wage Premia?

Our approach to estimating earnings elasticities from job switchers uses variation from both firm and occupation transitions. Appendix Figure A.5 splits the sample into firm and occupation switchers, showing that the earnings responses to lower taxes are similar in the two subsamples.<sup>27</sup> 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 taxes.

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

$$\log z_{it} = \alpha_i + \psi_{J(i,t)} + \mathbf{X}_{it}\boldsymbol{\beta} + \nu_{it}, \quad (19)$$

where  $\alpha_i$  is an individual fixed effect,  $\psi_{J(i,t)}$  is a firm fixed effect, and  $\mathbf{X}_{it}$  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 in 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 run the following regression in the sample of firm switchers:

$$\Delta \hat{\psi}_{J(i,t)} = \beta_0 + \sum_{b \neq b_0} \beta_b \cdot \mathbf{1}[y_{i0} \in b] + \mu_{it}. \quad (20)$$

The difference-in-differences coefficient  $\beta_b$  captures the effect of tax reform on the firm-specific earnings premia of firm switchers in income bin  $b$ . 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 15. It plots the changes in firm-specific earnings premia by

---

<sup>27</sup>Job switchers may change firm and occupation simultaneously. To retain statistical power, the subsamples in Figure A.5 include either all firm switches (even if occupation also changes) or all occupation switches (even if firm also changes). The results for within-occupation firm switches are similar to those shown in the figure, but the results for within-firm occupation switches are noisy as such switches represent a small fraction of the data.

income bin (the coefficients  $\hat{\beta}_b$ ) in different time intervals: 2006-08 (placebo), 2008-10, 2010-12, and 2012-14. In every time interval, these changes are close to zero and statistically insignificant at all income levels above and below the treatment threshold. In other words, the earnings responses of 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.

## 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 specific 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 precisely estimated, but modest, earnings elasticity of about 0.1. Our job switcher approach, on the other hand, gives an elasticity of 0.4-0.5. 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 (as reviewed by [Saez, Slemrod, and Giertz 2012](#)).

While we argue that the long-run elasticity is larger than typically estimated, our analysis does not support the extremely large elasticities implied by some macro calibrations. What are the policy implications of an elasticity of 0.4 as opposed to an elasticity of 0.1? As an example, consider the Laffer rate on top earners. This is determined by the classic formula  $\tau = 1 / (1 + \varepsilon a)$ , where  $\varepsilon$  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 0.1 implies  $\tau = 0.87$ , while an elasticity of 0.4 implies  $\tau = 0.62$ . Therefore, the long-run elasticity we estimate has quantitatively large policy implications.<sup>28</sup> The implications are even larger in countries with more compressed earnings distributions because their larger Pareto parameter magnifies the impact of the elasticity.

---

<sup>28</sup>Still, even at an elasticity of 0.4, the current top marginal tax rate in the US is well below the Laffer point, which is also the optimal top tax rate absent any non-tax externalities.

## 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](#), [29](#)
- 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. [5](#)
- CABALLERO, RICARDO J., AND EDUARDO M.R.A. ENGEL (1991): "Dynamic (S,s) Economies," NBER Working Paper Working No. 3734. [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](#), [31](#)
- 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](#)
- HARRIS, MILTON, AND BENGT HOLMSTRÖM (1982): “A Theory of Wage Dynamics,” *The Review of Economic Studies*, 49(3), 315–333. [6](#)
- 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. [14](#)
- 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](#), [28](#)
- 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](#), [21](#), [22](#)
- 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](#)
- (2023): “The EITC and the Extensive Margin: A Reappraisal,” NBER Working Paper No. 26405. [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. [14](#)
- 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](#)
- 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](#), [31](#)
- (2010): "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2(3), 180–212. [7](#)
- 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. [5](#), [10](#), [21](#), [22](#), [30](#)
- SCHEUER, FLORIAN, AND IVÁN WERNING (2017): "The Taxation of Superstars," *The Quarterly Journal of Economics*, 132(1), 211–270. [6](#)
- SIGAARD, HANS SCHYTTE (2022): "Labor Supply Responsiveness to Tax Reforms," Working Paper. [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](#)



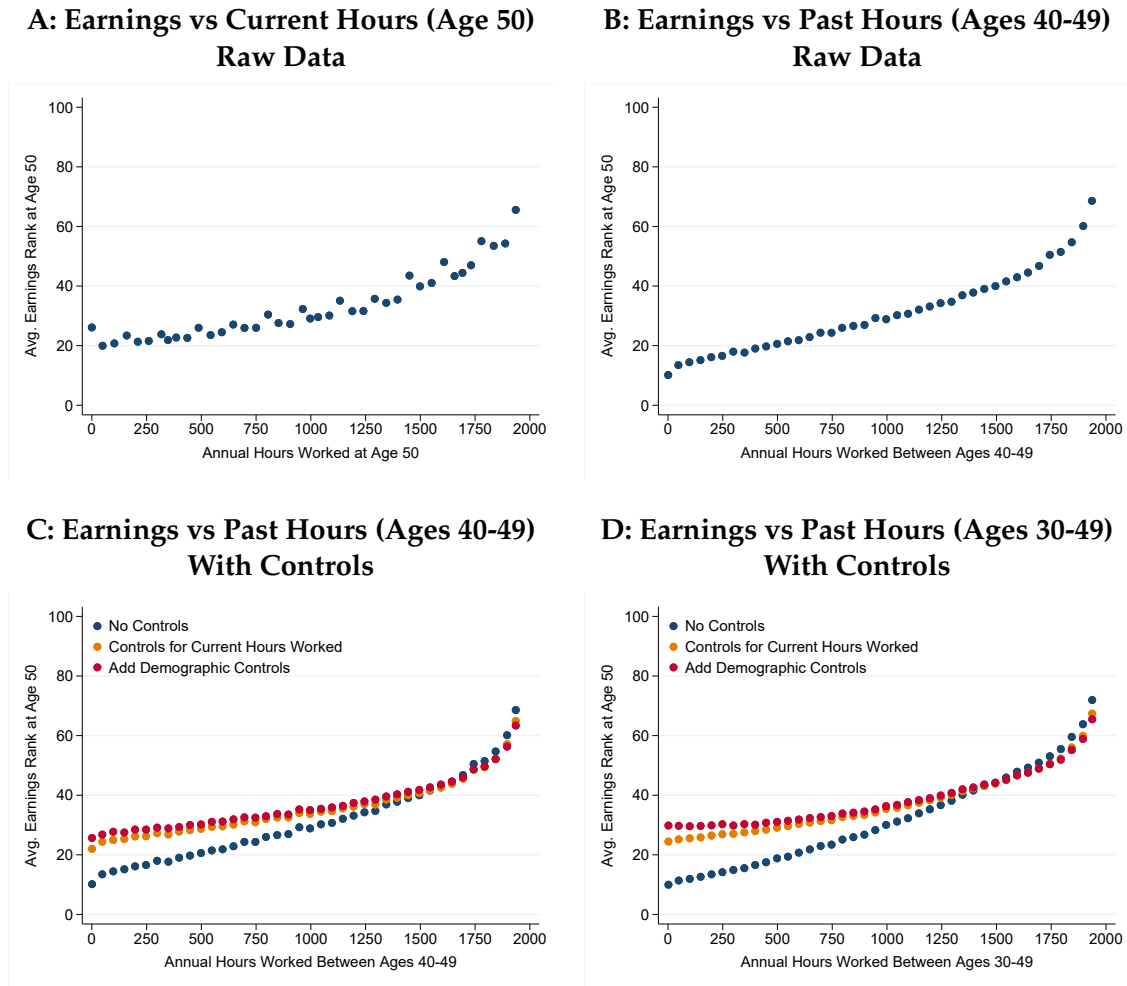
TORTAROLO, DARIO, GUILLERMO CRUCES, AND VICTORIA CASTILLO (2020): “It takes two to tango: Labor responses to an income tax holiday in Argentina,” Working Paper. [5](#)

**TABLE 1: ELASTICITY ESTIMATES**  
STANDARD VS DYNAMIC APPROACH

	$\Delta \text{Log Earnings}$		Elasticity	
	ITT	TOT	ITT	TOT
<b>Standard Approach: All Workers</b>				
<i>Time Period:</i>				
2008-10	0.011 (0.001)	0.016 (0.002)	0.06 (0.007)	0.10 (0.011)
2008-11	0.007 (0.001)	0.011 (0.002)	0.04 (0.007)	0.06 (0.012)
2008-12	0.001 (0.001)	0.002 (0.002)	0.01 (0.008)	0.01 (0.013)
2008-13	-0.009 (0.001)	-0.016 (0.002)	-0.05 (0.008)	-0.09 (0.014)
2008-14	-0.013 (0.001)	-0.022 (0.003)	-0.07 (0.008)	-0.13 (0.015)
<b>Dynamic Approach: Movers vs Stayers</b>				
<i>Time Period:</i>				
2008-10	0.057 (0.003)	0.086 (0.004)	0.33 (0.015)	<b>0.49 (0.022)</b>
2008-11	0.053 (0.003)	0.083 (0.004)	0.31 (0.015)	<b>0.48 (0.023)</b>
2008-12	0.052 (0.003)	0.085 (0.004)	0.30 (0.016)	<b>0.49 (0.025)</b>
2008-13	0.047 (0.003)	0.081 (0.005)	0.27 (0.016)	<b>0.47 (0.028)</b>
2008-14	0.045 (0.003)	0.077 (0.005)	0.26 (0.017)	<b>0.44 (0.030)</b>

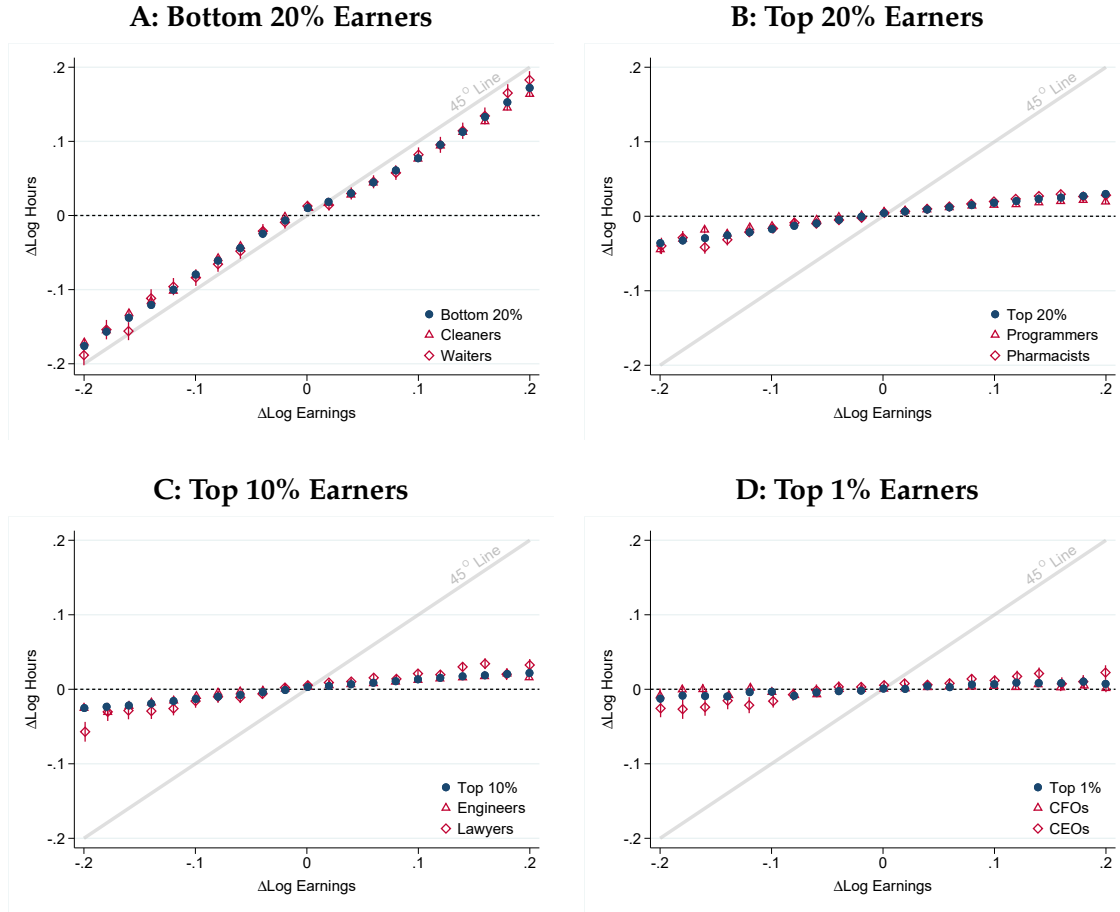
Notes: This table reports estimates of log-earnings responses and elasticities with respect to  $1 - \tau$  based on the standard approach (top panel) and the dynamic approach (bottom panel). The standard approach estimates the average effect of taxes on all workers (applying a triple-differences specification to the 2009 reform) while the dynamic approach estimates the effect on job movers relative to job stayers (applying a quadruple-differences specification to the 2009 reform), as described in section 5.1. The table shows both intention-to-treat (ITT) and treatment-on-the-treated (TOT) effects. The ITT estimates assign treatment status based on pre-reform tax bracket, while the TOT estimates assign treatment status based on actual tax bracket, instrumented using the pre-reform bracket. To ensure that pre-reform bracket assignment is a strong predictor of actual bracket assignment, the sample excludes taxpayers located within 30,000 DKK of the treatment threshold ex ante. The table provides estimates of behavioral responses over varying time intervals: 2008-10, 2008-11, 2008-12, 2008-13, and 2008-14. The key estimates in the table are the TOT elasticities based on the dynamic approach. These estimates are large and very stable over time. Robust standard errors are provided in parentheses.

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



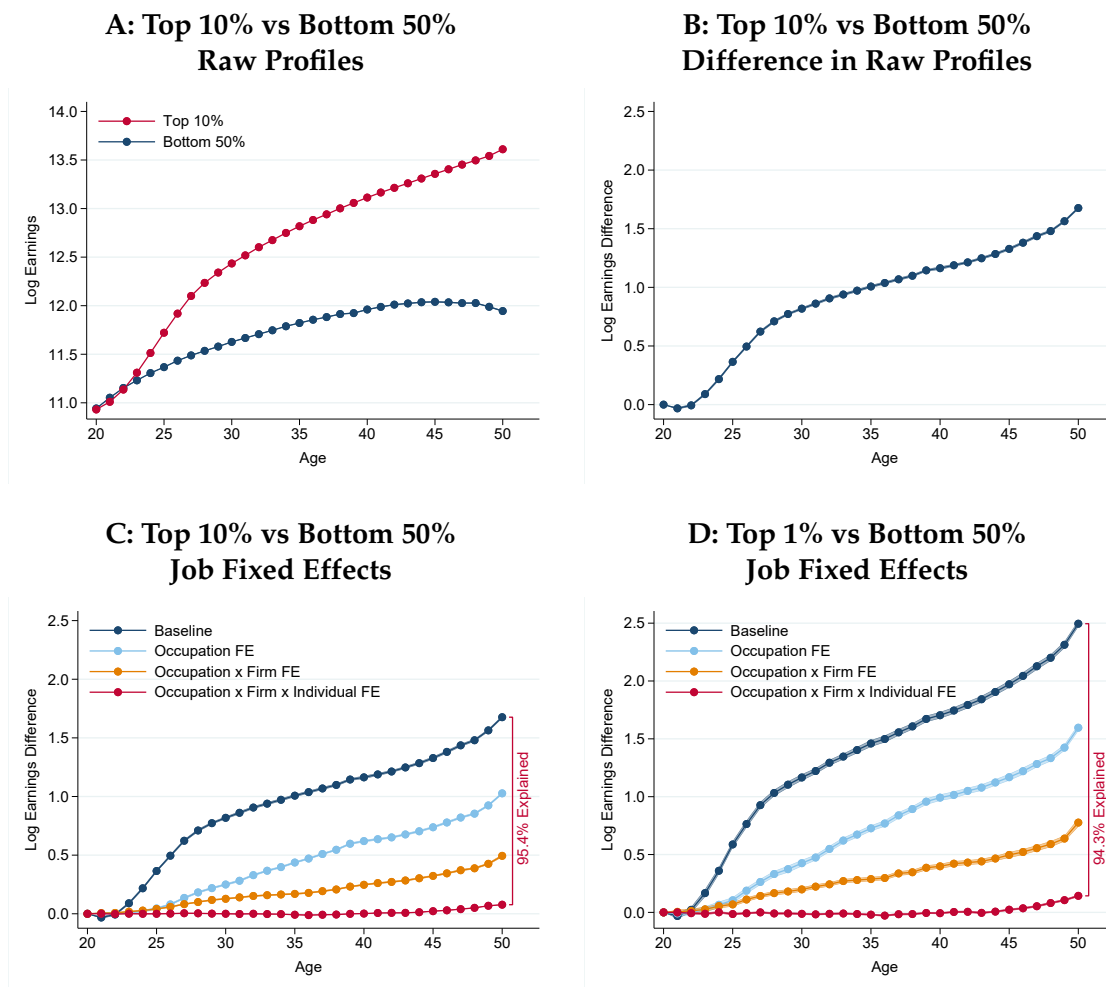
Notes: This figure shows that past hours worked predict current earnings. The figure is based on a balanced panel of workers observed between the ages of 20 and 50. Each panel plots the relationship between average earnings rank at age 50 and hours worked at different ages: current hours in Panel A and 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 and demographic characteristics (dummies for gender, children, marital status, and education level). The graphs include 95% confidence intervals based on standard errors clustered at the individual level, but these are hardly visible.

**FIGURE 2: CONTEMPORANEOUS HOURS AND EARNINGS CHANGES ARE UNRELATED AT THE TOP, BUT NOT AT THE BOTTOM**



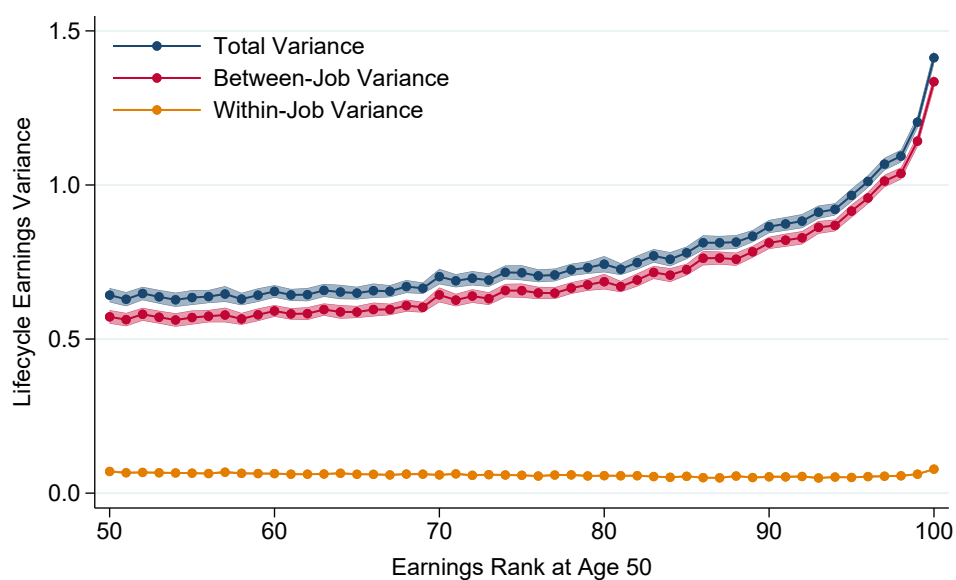
Notes: This figure shows the contemporaneous relationship between hours and earnings changes at the intensive margin in different segments of the earnings distribution. It plots changes in log hours against changes in log earnings in the bottom 20%, the top 20%, the top 10%, and the top 1% of the distribution. 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 3: LIFECYCLE PROFILES IN EARNINGS ARE DRIVEN BY DISCRETE JOB SWITCHES**



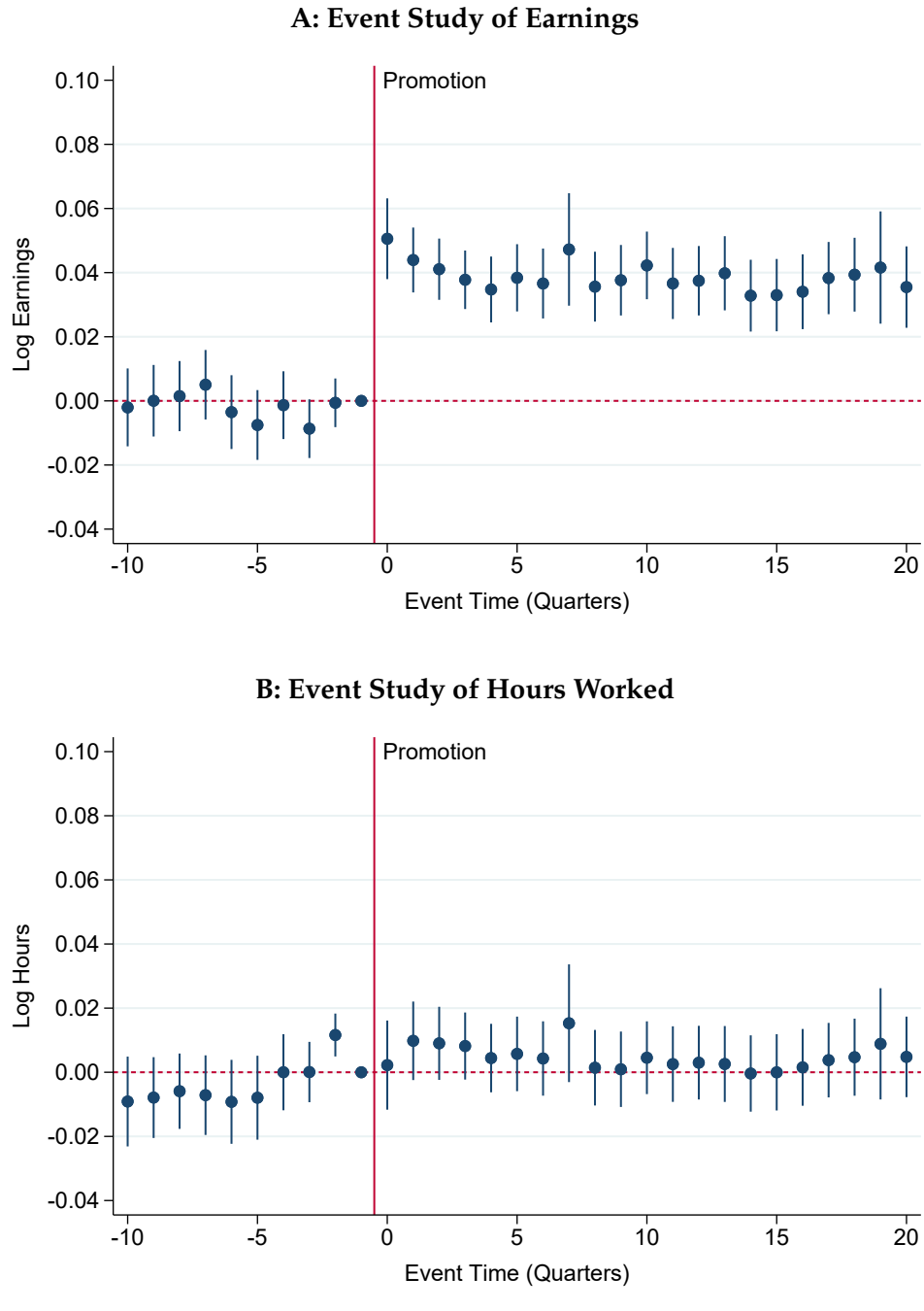
Notes: This figure shows that the lifecycle profile of earnings for workers who reach the top of the distribution is driven by job switches, defined as transitions between occupation $\times$ firm cells. The figure is based on a balanced panel of workers observed between the ages of 20 and 50, plotting the earnings profiles of workers observed in different quantiles of the distribution at age 50. Panel A plots the raw profiles of 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 $\times$ firm fixed effects (orange), and net of occupation $\times$ firm $\times$ individual fixed effects (red). Within individual, job fixed effects explain about 95% 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 results are very similar. The shaded areas (not always visible) represent 95% confidence intervals based on standard errors clustered at the individual level.

**FIGURE 4: BETWEEN-JOB VS WITHIN-JOB VARIATION IN EARNINGS**  
 VARIANCE DECOMPOSITION BY EARNINGS PERCENTILE AT AGE 50



Notes: This figure decomposes the variance of lifecycle earnings into between-job variance and within-job variance using equation (15). As in the preceding figure, jobs are measured as occupation $\times$ firm cells and the estimation 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 $\times$ firm cells. The shaded areas depict 95% confidence intervals based on standard errors clustered at the individual level.

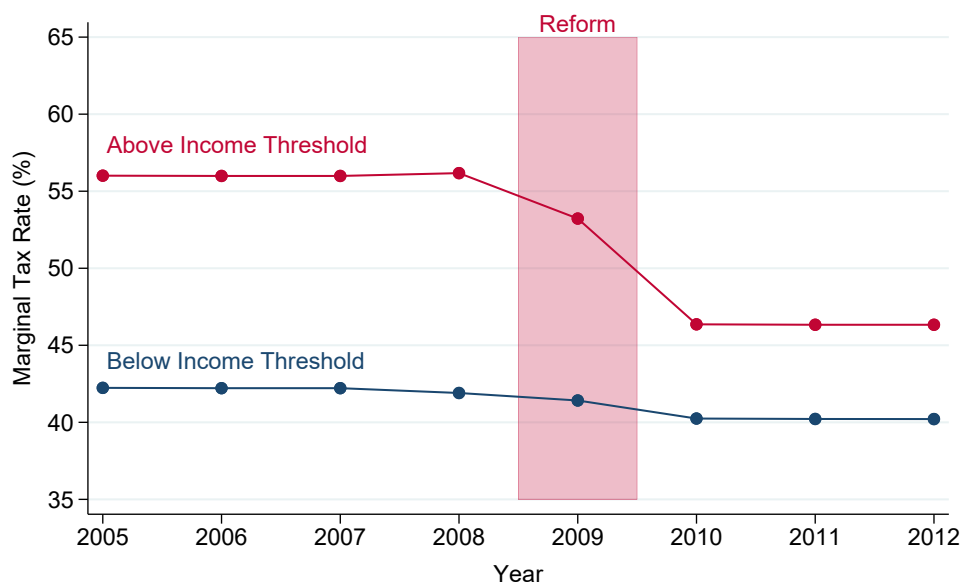
**FIGURE 5: EARNINGS JUMP DISCRETELY AT PROMOTIONS, WITH NO CHANGE IN HOURS**



Notes: This figure presents event studies of promotions using specification (16). 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, with no effect on hours worked. The error bars depict 95% confidence intervals based on standard errors clustered at the individual level.

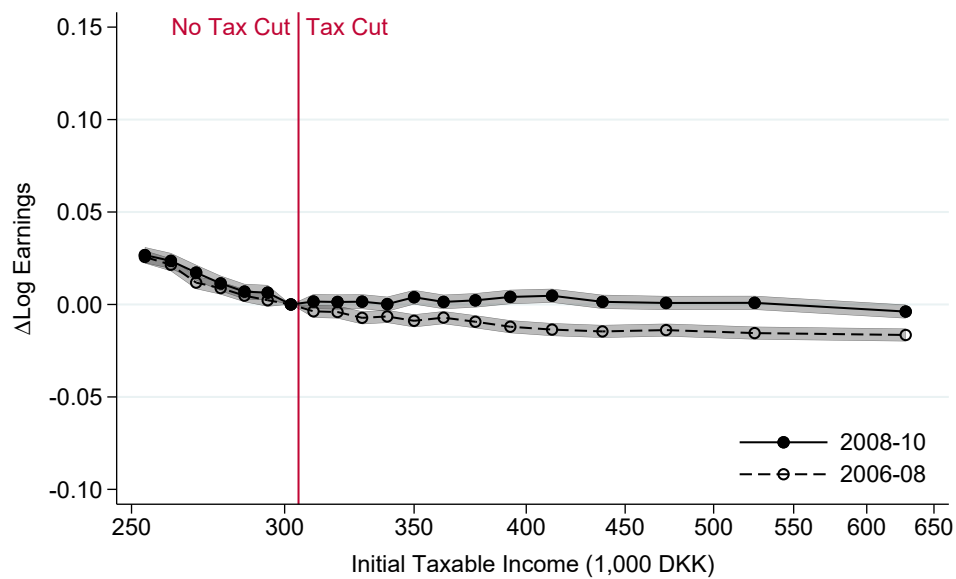


**FIGURE 6: 2009 TAX REFORM IN DENMARK**



Notes: This figure illustrates the tax variation created by the 2009 reform in Denmark. The reform reduced marginal tax rates above an income threshold, leaving marginal tax rates below the threshold roughly unchanged. The threshold that separates treatments and controls is located at a taxable income of 306,000 Danish Kroner (2008 prices), corresponding roughly to the 70th percentile of the distribution. The figure shows marginal tax rates over time for taxpayers above the threshold (red series) and below the threshold (blue series). The reform-induced reduction in the top marginal tax rate was almost 10 percentage points.

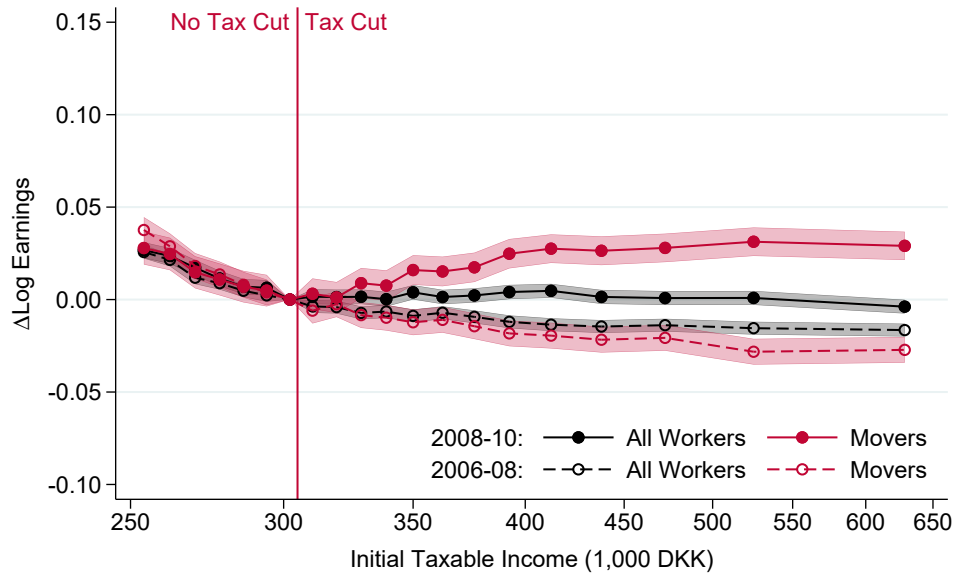
**FIGURE 7: IMPACT OF TAX REFORM ON EARNINGS**  
ALL WORKERS



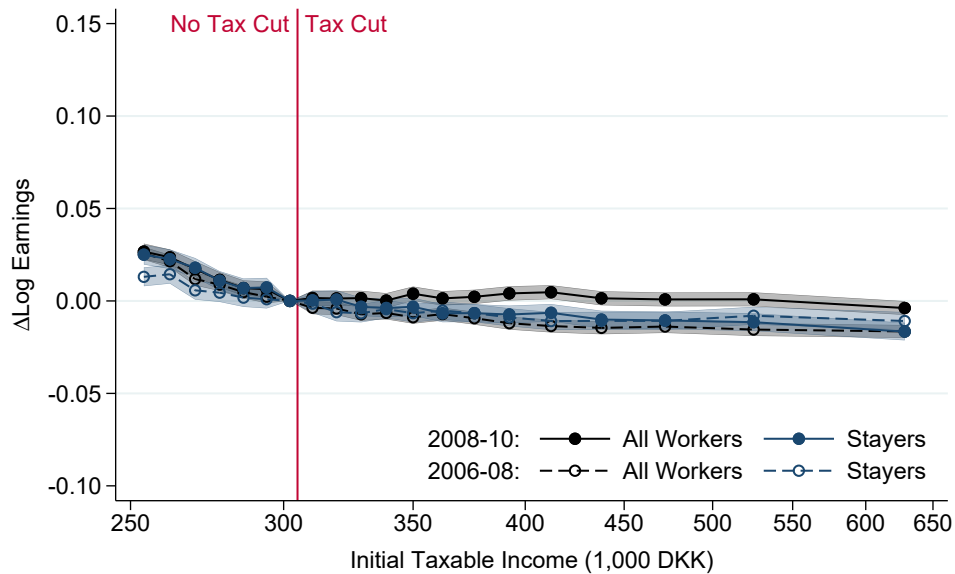
Notes: This figure shows the impact of the 2009 tax reform on earnings in the full sample of workers. It plots changes in log earnings between 2008-10 (post-reform) and between 2006-08 (pre-reform placebo) by income bin, omitting the bin just below the treatment threshold depicted by the vertical line. The 2008-10 and 2006-08 series track each other closely below the threshold and diverge above the threshold, consistent with earnings responses to the reduction in marginal tax rates. The shaded areas show 95% confidence intervals based on robust standard errors.

**FIGURE 8: IMPACT OF TAX REFORM ON EARNINGS**  
SWITCHERS VS NON-SWITCHERS

**A: Switchers**

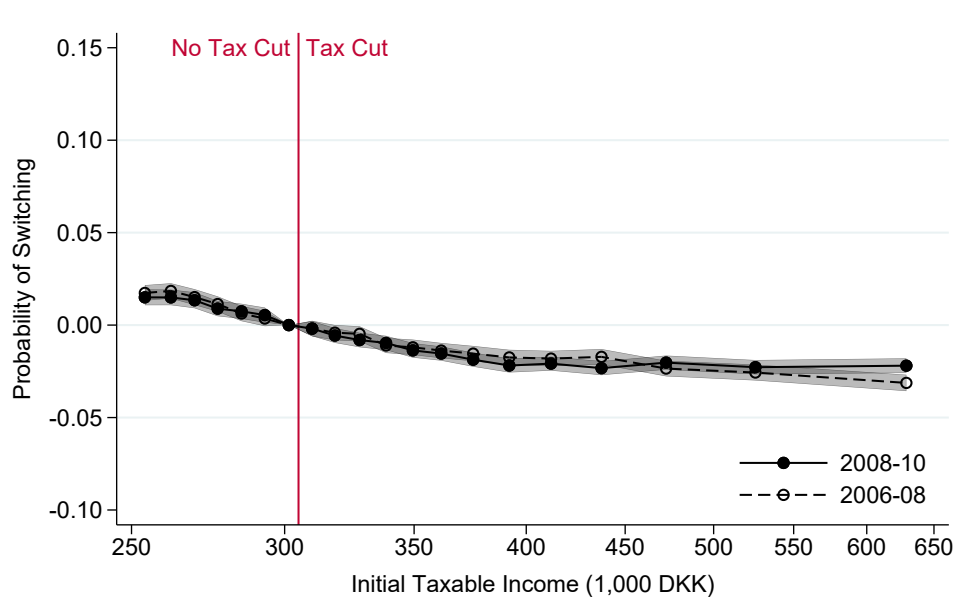


**B: Non-Switchers**



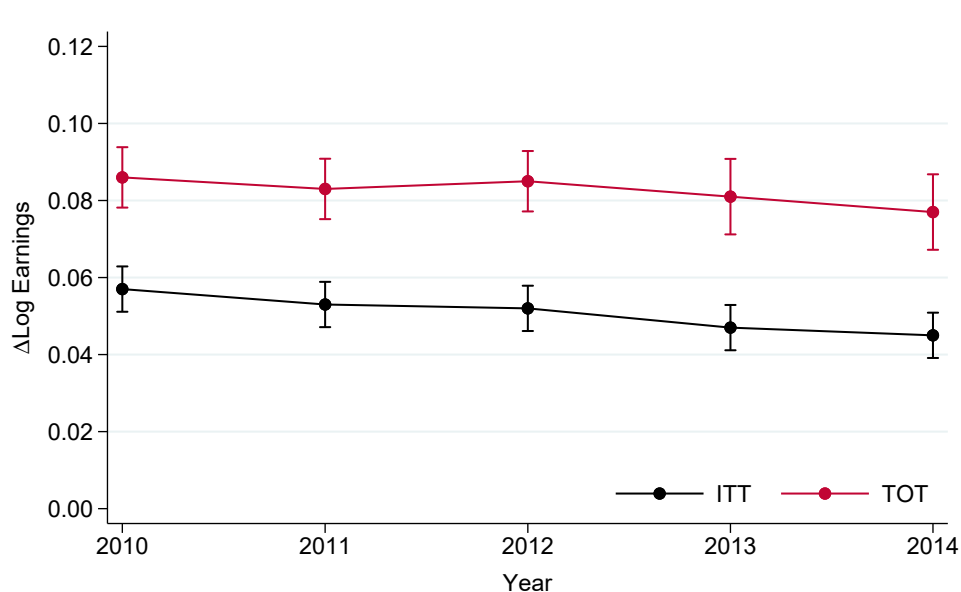
Notes: This figure shows the impact of the 2009 tax reform on earnings for job switchers (Panel A) and job non-switchers (Panel B). Switchers are those who move between occupation×firm cells, while non-switchers are those who stay within occupation×firm cells. Each panel plots changes in log earnings between 2008-10 (post-reform) and between 2006-08 (pre-reform placebo) by income bin, omitting the bin just below the treatment threshold depicted by the vertical line. For job movers, the 2008-10 and 2006-08 series track each other closely below the threshold and diverge sharply above the threshold. The earnings responses among movers are much larger than in the full sample. For job stayers, the two series track each other both below and above the threshold, implying that there are no earnings responses among these workers. The shaded areas show 95% confidence intervals based on robust standard errors.

**FIGURE 9: IMPACT OF TAX REFORM ON SWITCHING PROBABILITY**



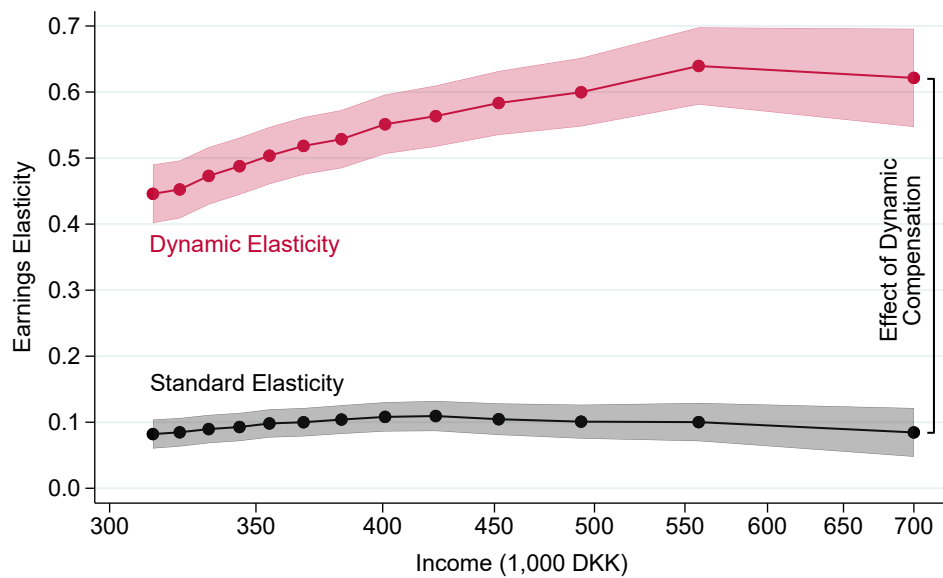
Notes: This figure shows the impact of the 2009 tax reform on the probability of switching jobs, defined as moving between occupation×firm cells. The figure plots the switching probability between 2008-10 (post-reform) and 2006-08 (pre-reform) by income bin, omitting the bin just below the treatment threshold depicted by the vertical line. The 2008-10 and 2006-08 series track each other closely both below and above the threshold, implying that the reform has no effect on the probability of switching jobs. The shaded areas show 95% confidence intervals based on robust standard errors.

**FIGURE 10: IMPACT OF TAX REFORM ON EARNINGS OVER TIME**  
AVERAGE RESPONSE BY JOB SWITCHERS



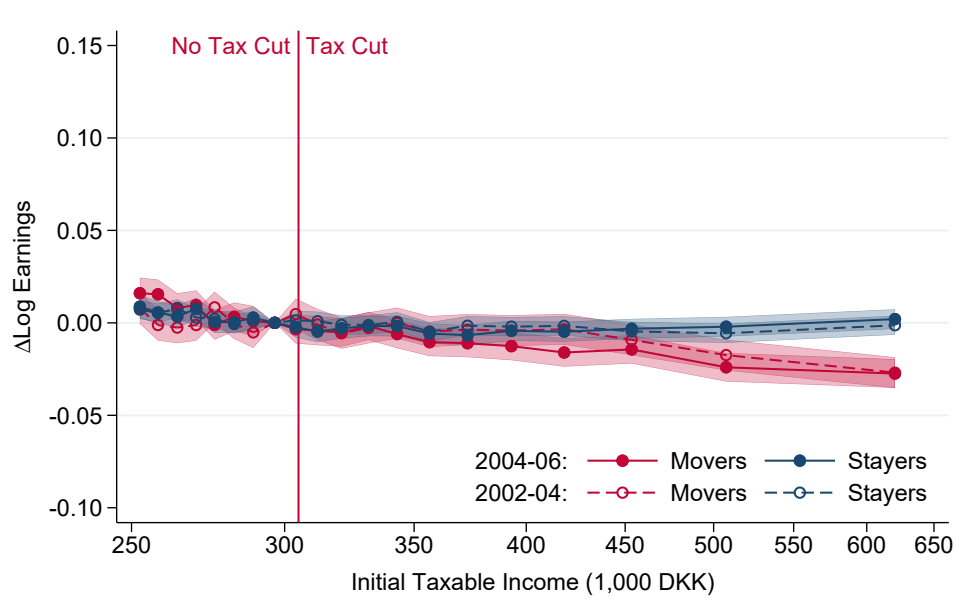
Notes: This figure shows the impact of the 2009 tax reform on earnings over time. The focus is on the average impact on job switchers in the following time intervals: 2008-10, 2008-11, 2008-12, 2008-13, and 2008-14. All estimates are based on the quadruple-differences specification described in section 5.1. The figure shows both intention-to-treat (ITT) and treatment-on-the-treated (TOT) effects. The ITT estimates assign treatment status based on pre-reform tax bracket, while the TOT estimates assign treatment status based on actual tax bracket, instrumented using the pre-reform bracket. To ensure that pre-reform bracket assignment is a strong predictor of actual bracket assignment, the sample excludes taxpayers located within 30,000 DKK of the treatment threshold ex ante. The figure shows that the earnings responses of job switchers are stable over time. This is especially true of the TOT estimates, which correct for attenuation bias due to taxpayers changing bracket location over time. The error bars depict 95% confidence intervals based on robust standard errors.

**FIGURE 11: DYNAMIC VS STANDARD ELASTICITY BY INCOME LEVEL**



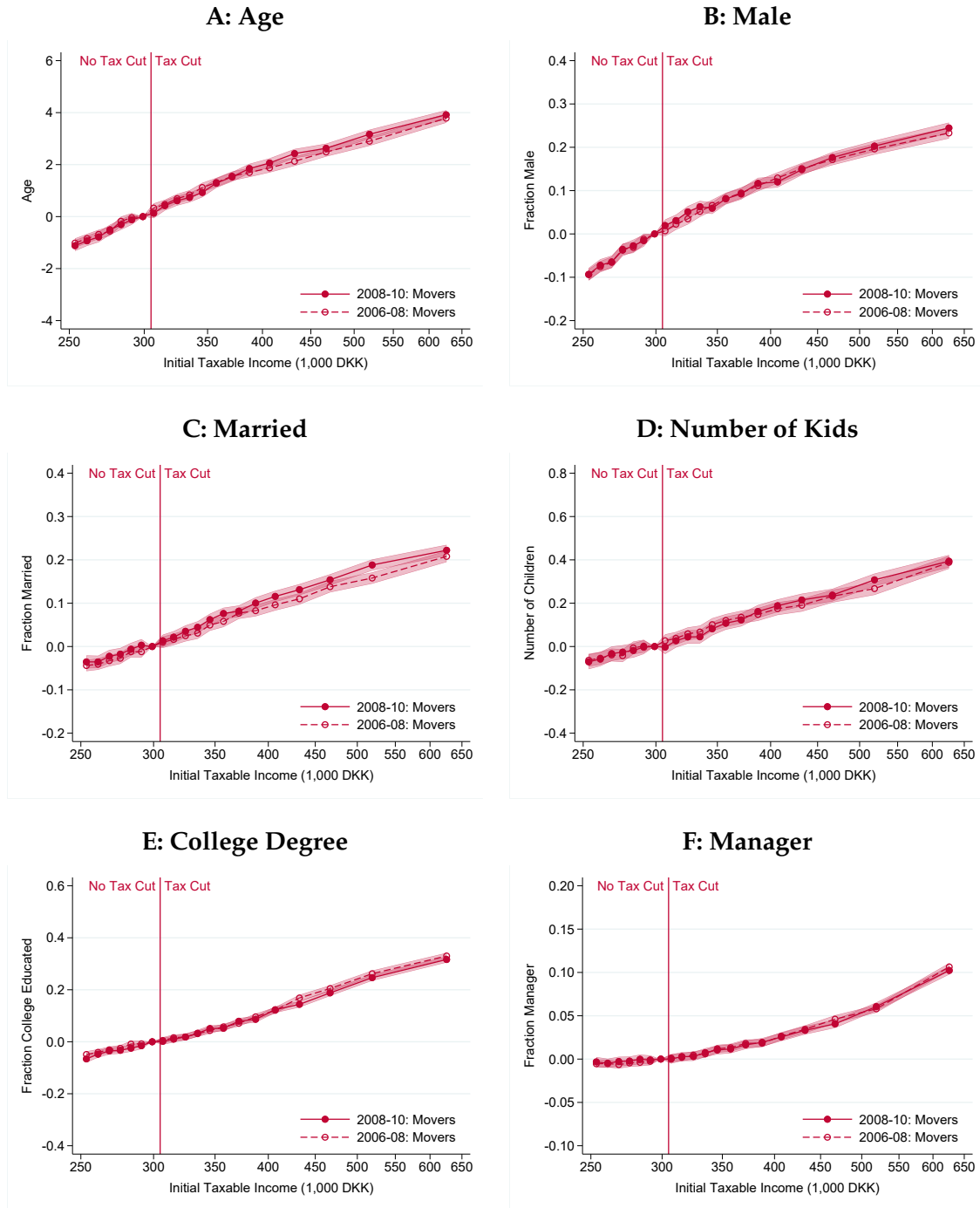
Notes: This figure shows earnings elasticities by income level based on the standard approach (black series) and the dynamic approach (red series). The standard approach captures the average effect on all workers (estimated using a triple-differences specification) while the dynamic approach captures the effect on job movers relative to job stayers (estimated using a quadruple-differences specification), as described in section 5.1. For each approach, the figure shows elasticities between 2008-10 above different income cutoffs, depicted on the x-axis. The standard elasticity is small and virtually constant in income. The dynamic elasticity is large and increasing in income. The gap between the two — the effect of dynamic compensation realized at the point of switching — is therefore increasing in the level of income. The shaded areas show 95% confidence intervals based on robust standard errors.

**FIGURE 12: ZERO IMPACT OF PLACEBO REFORM**  
SWITCHERS VS NON-SWITCHERS



Notes: This figure shows the impact of a placebo reform on job movers and job stayers. It plots changes in log earnings between 2004-06 (after placebo reform) and 2002-04 (before placebo reform) by income bin, omitting the bin just below the treatment threshold depicted by the vertical line. The post- and pre-reform series track each other closely across all income levels for both movers and stayers. Hence, the placebo reform has zero impact in both samples, consistent with causal identification in our main specification. The shaded areas show 95% confidence intervals based on robust standard errors.

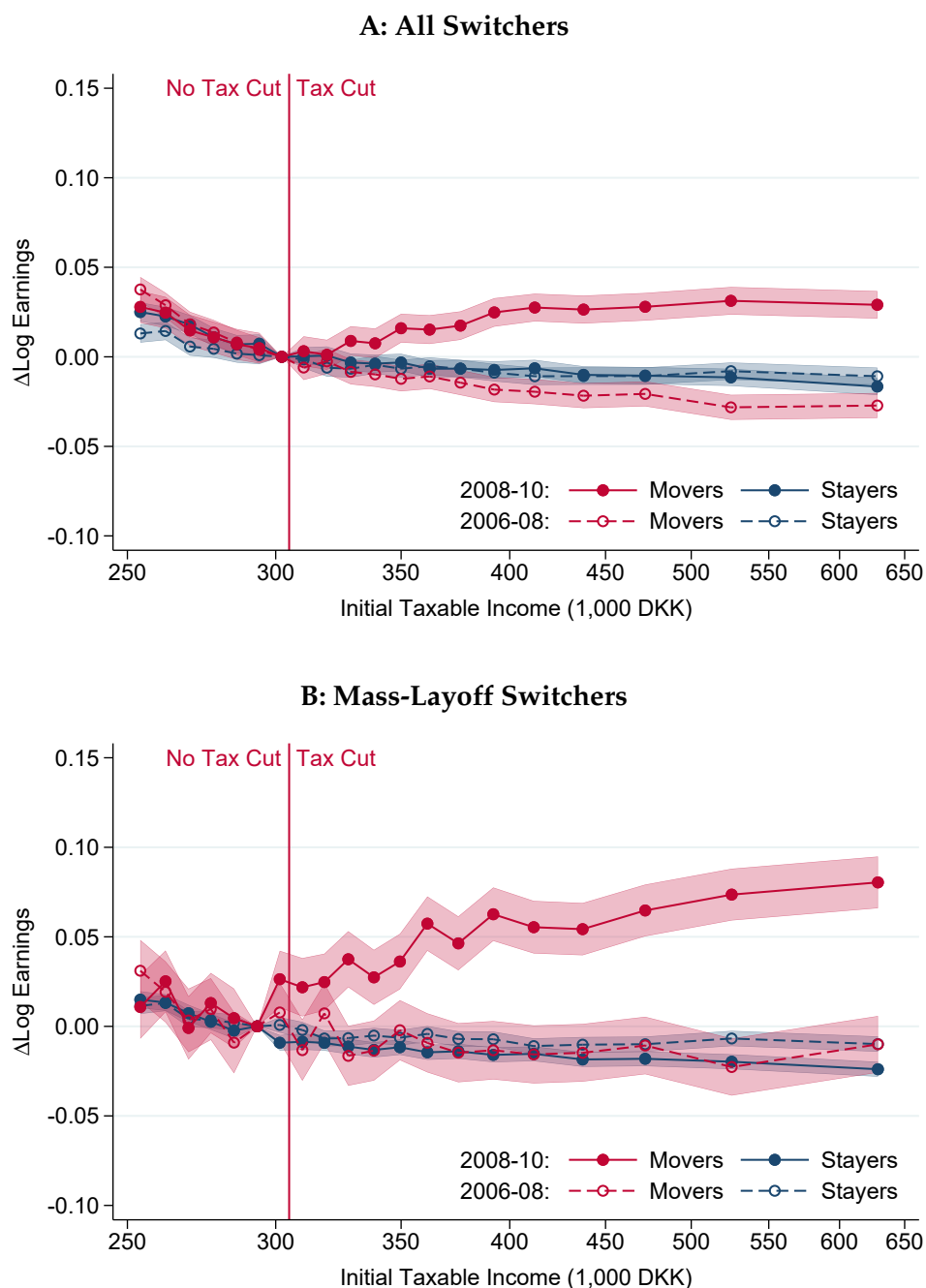
**FIGURE 13: IS SWITCHING SELECTED?**  
IMPACT OF TAX REFORM ON SWITCHER CHARACTERISTICS



Notes: This figure investigates if job switchers are selected in a way that poses a threat to our empirical design. This is done by treating demographic characteristics as dependent variables in specification (17), asking if any of them respond to the reform. Each panel of the figure plots a demographic variable by income bin for workers who switch jobs between 2008-10 (after the reform) and between 2006-08 (before the reform). Six different variables are shown: age, fraction male, fraction married, number of children, fraction college educated, and fraction being manager. All of these variables are increasing in income — and hence larger among treatments than among controls — but the relationship is virtually identical before and after the reform. This implies that the reform has no impact on any observable characteristics of job switchers, suggesting that our estimates are not biased by selection of job switchers.

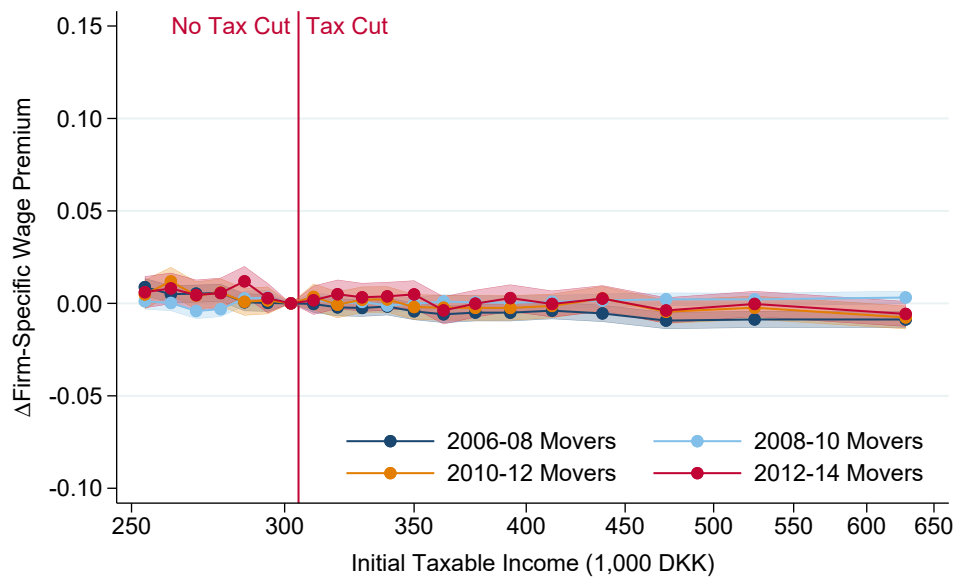


**FIGURE 14: IS SWITCHING SELECTED?**  
ALL SWITCHERS VS MASS-LAYOFF SWITCHERS



Notes: This figure shows the impact of the 2009 tax reform on earnings in the full sample of switchers (Panel A) and in the 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% and has at least 20 employees ex ante. Each panel plots changes in log earnings between 2008-10 (post-reform) and 2006-08 (pre-reform placebo) by income bin. The earnings responses in the mass-layoff sample are qualitatively similar, but somewhat larger in magnitude, than the responses in the full sample of switchers. This suggests that the large switcher elasticities estimated above are not driven by selection into job switching. The shaded areas show 95% confidence intervals based on robust standard errors.

**FIGURE 15: IMPACT OF TAX REFORM ON FIRM-LEVEL WAGE PREMIA OF SWITCHERS**



Notes: This figure investigates if the earnings responses to lower taxes are mediated by firm-level wage premia. The estimation of firm-level wage premia is based on the AKM model in equation (19). Using the estimated wage premia as outcomes, the figure plots coefficients from the difference-in-differences specification (20) run on the sample of firm switchers. It plots changes in the firm-specific wage premia of firm switchers by income bin over different time intervals: 2006-08 (placebo), 2008-10, 2010-12, and 2012-14. In every time interval, the changes in firm-specific wage premia are close to zero and statistically insignificant at all income levels above and below the treatment threshold. 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.

## **Online Appendix**

### **A Supplementary Figures and Tables**

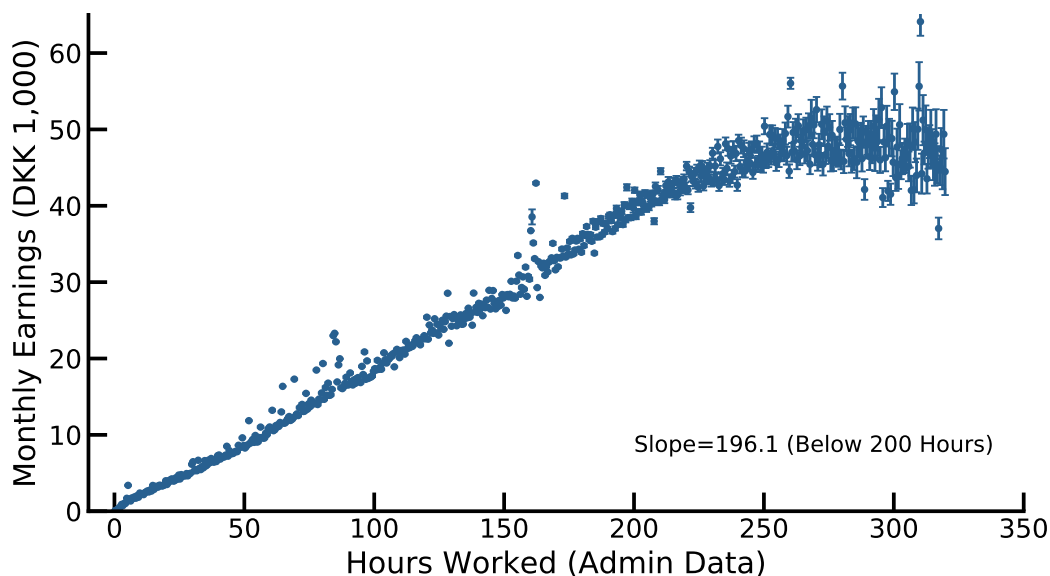
**TABLE A.1: OCCUPATIONS OF TOP AND BOTTOM EARNERS**  
JOB TITLES FOR 45-50 YEAR OLDS

	Occupation	Mean Income (DKK 1,000)	
		All Firms	Top 1% Firms
Top 10	Managers, Business Services	1,265.1	2,945.2
	Managers, Business Strategy	1,019.4	1,097.0
	Branch Managers, Financial and Insurance Services	983.5	967.8
	Managing Director and Chief Executives	964.4	1,547.7
	Lawyers	937.2	1,180.0
	Securities and Finance Dealers and Brokers	930.6	951.9
	Specialist Medical Doctors	884.0	818.5
	Aircraft Pilots	816.5	842.5
	Generalist Medical Doctors	811.2	842.5
	Senior Government Officials	797.4	814.4
Bottom 10	Nursing and Midwifery Associate Professionals	217.1	225.8
	Construction Workers	215.9	-
	Transport and Storage Workers	211.3	247.6
	Livestock Farm Workers	208.3	239.5
	Sports and Fitness Workers	207.2	237.6
	Messengers, Package Deliverers and Luggage Porters	203.9	192.4
	Cleaners and Helpers	199.8	233.4
	Manufacturing Workers	199.6	200.6
	Unskilled Workers in Mining, Manufacturing, etc.	194.1	-
	Street and Market Sales Persons	170.7	-

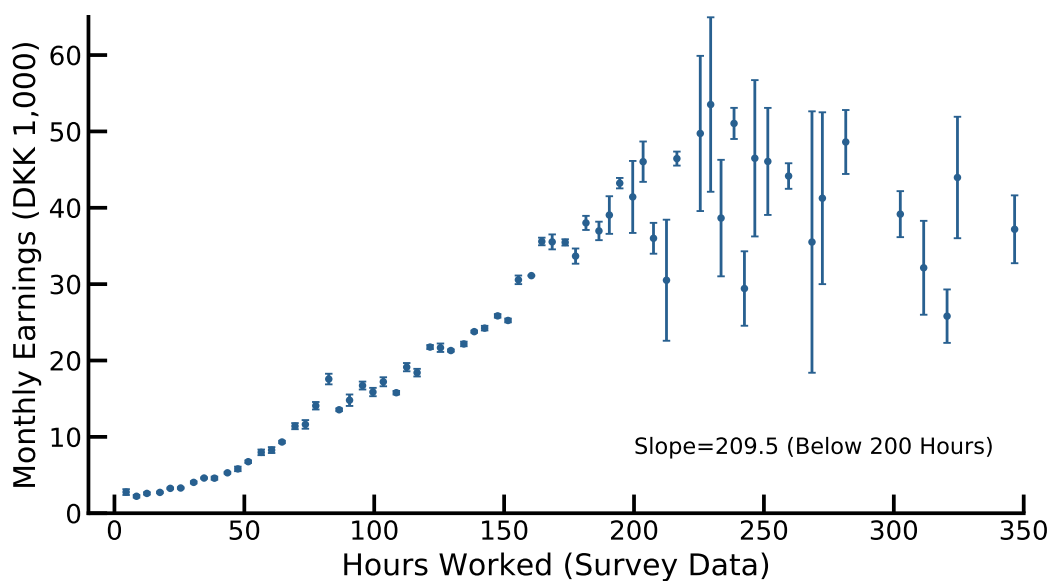
Notes: This table shows the highest-paying occupations (top panel) and the lowest-paying occupations (bottom panel) for workers aged 45-50. The occupation classification is based on 6-digit (D)ISCO codes as described in section 3, ranked by average wage earnings. For each occupation, the table reports average earnings across all firms and across the top 1% largest firms (in terms of number of employees).

**FIGURE A.1: VALIDATION OF ADMINISTRATIVE HOURS WORKED MEASURE**  
ADMINISTRATIVE DATA VS SURVEY DATA

**A: Administrative Data (Pension Measure of Hours Worked)**



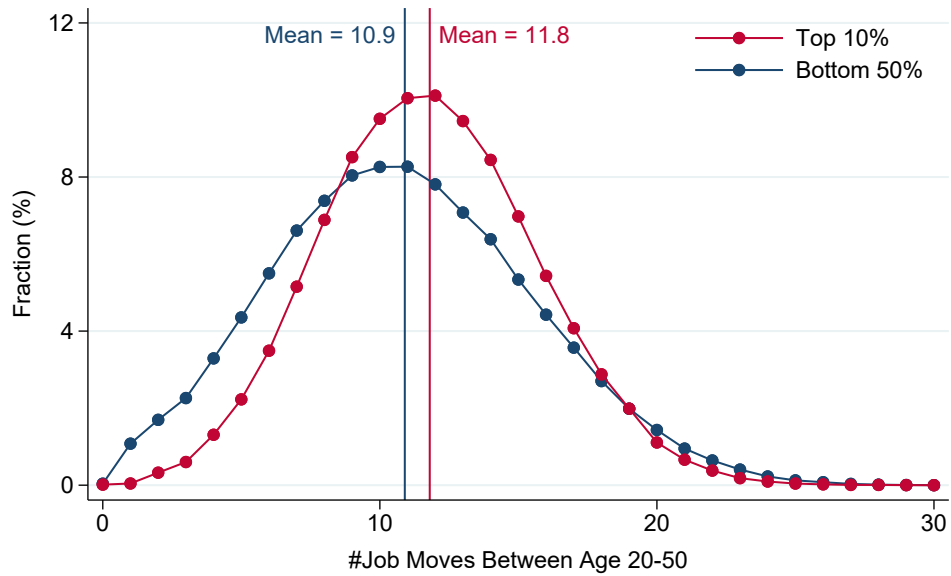
**B: Survey Data (Self-Reported Measure of Hours Worked)**



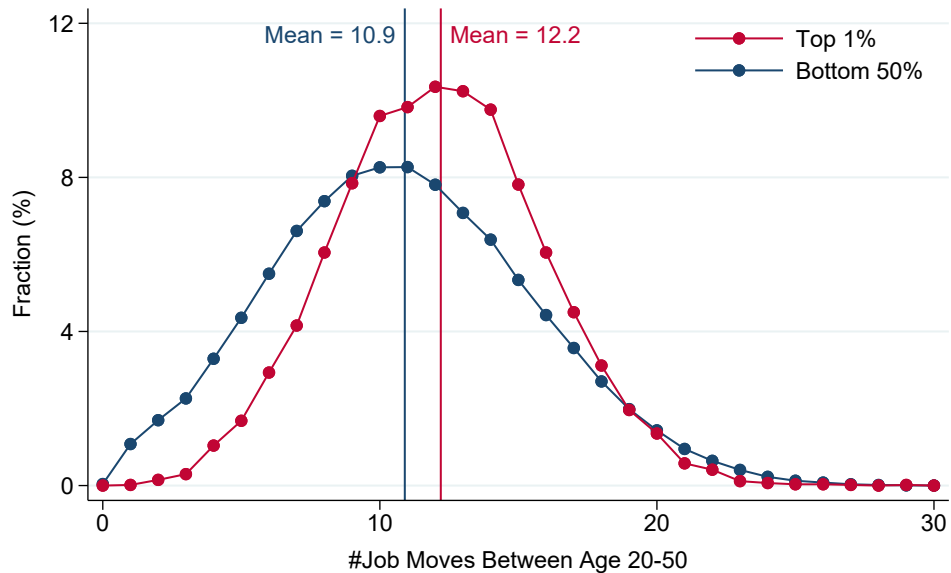
Notes: This figure validates the administrative measure of hours worked — the pension measure described in section 3 — against a survey measure of hours worked. The survey measure is based on a question about actual, uncapped hours taken from the Danish component of the EU Labour Force Survey. The figure 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. But 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: DISTRIBUTION OF NUMBER OF SWITCHES**  
TOP VS BOTTOM EARNERS BETWEEN AGES 20-50

**A: Top 10% vs Bottom 50%**



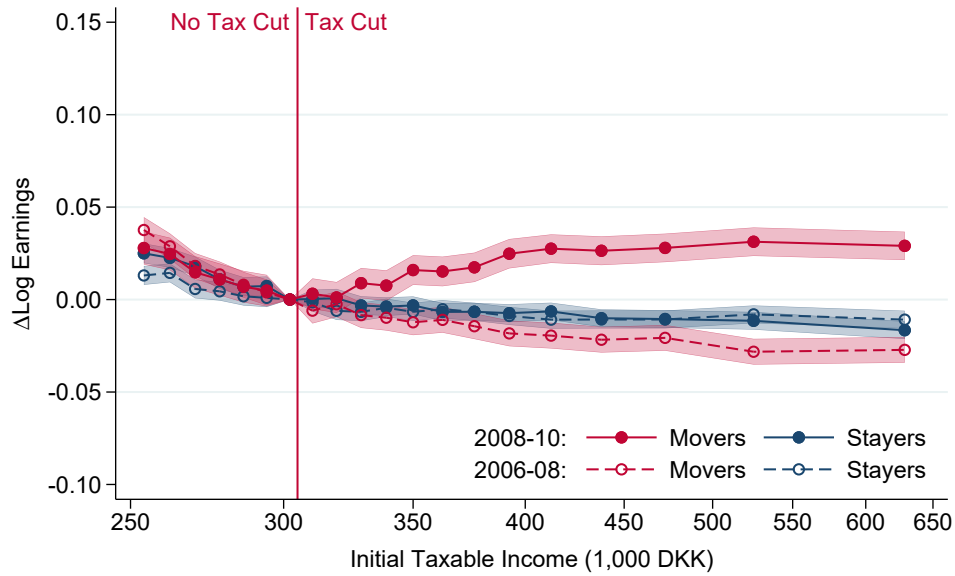
**B: Top 1% vs Bottom 50%**



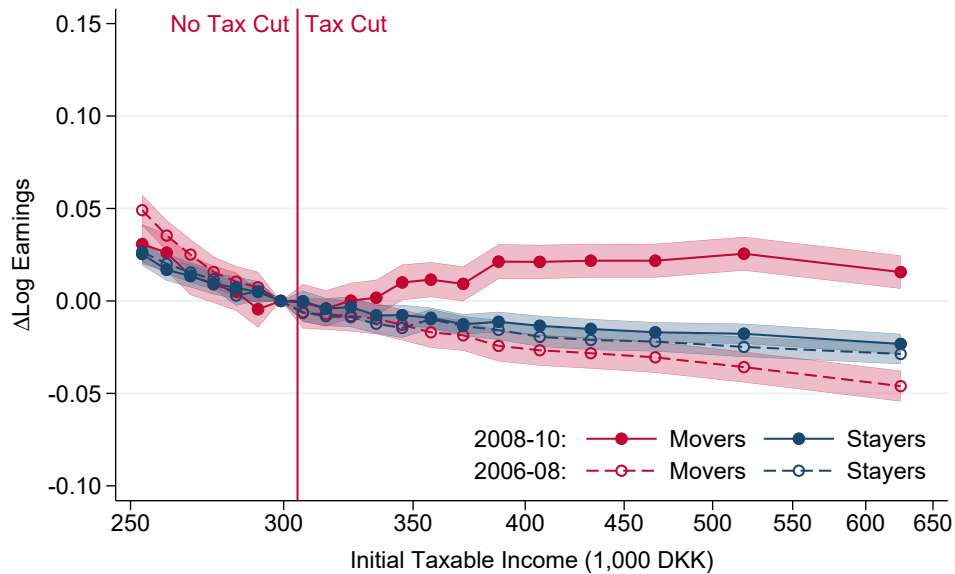
Notes: This figure shows the distribution of the number of job switches between ages 20-50 for top and bottom earners. The figure is based on a balanced panel of workers observed between ages 20-50, splitting the sample by their earnings percentile at age 50. Panel A compares top-10% and bottom-50% earners, while Panel B compares top-1% and bottom-50% earners. The distribution of number of switches is broadly similar among top and bottom earners. The average number of switches is about 12 at the top (11 at the bottom), corresponding to roughly one switch every three years.

**FIGURE A.3: IMPACT OF TAX REFORM ON EARNINGS**  
CONTROLLING FOR SWITCHER CHARACTERISTICS

**A: No Controls**

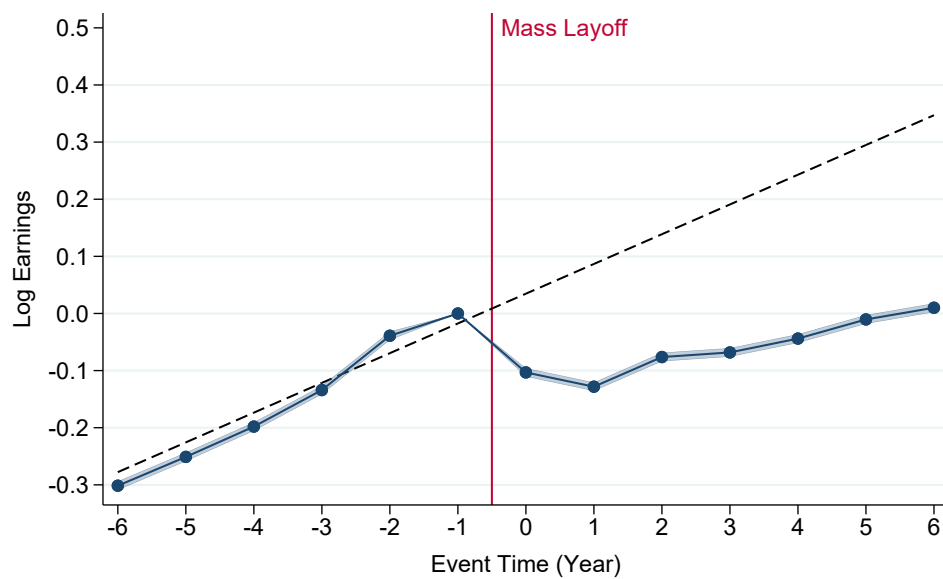


**B: Controls**



Notes: This figure shows the impact of the 2009 tax reform on earnings for movers and stayers in our baseline specification without demographic controls (Panel A) and in a specification with demographic controls (Panel B). The controls include dummies for education level (8 categories), gender (binary), children (binary), and marital status (7 categories). The figure is otherwise constructed in the same way as previous figures. It plots changes in log earnings between 2008-10 (post-reform) and between 2006-08 (pre-reform placebo) by income bin, omitting the bin just below the treatment threshold depicted by the vertical line. The empirical patterns are very similar in the two panels. This is consistent with the results in Figure 13, which shows that job movers are not selected on observables given our empirical design. The shaded areas show 95% confidence intervals based on robust standard errors.

**FIGURE A.4: IMPACT OF MASS LAYOFF ON EARNINGS**

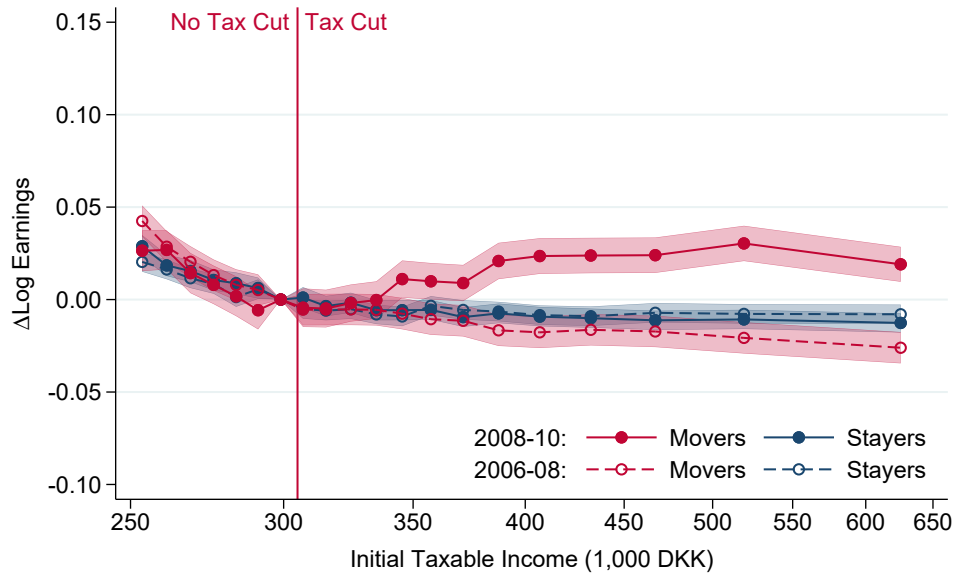


Notes: This figure presents an event study of the effect of mass layoffs on earnings. Mass layoffs are defined as layoffs in which firms with at least 20 employees reduce their workforce by at least 30% in a single year. The figure shows log earnings by event time (blue series) compared to a linear time 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.

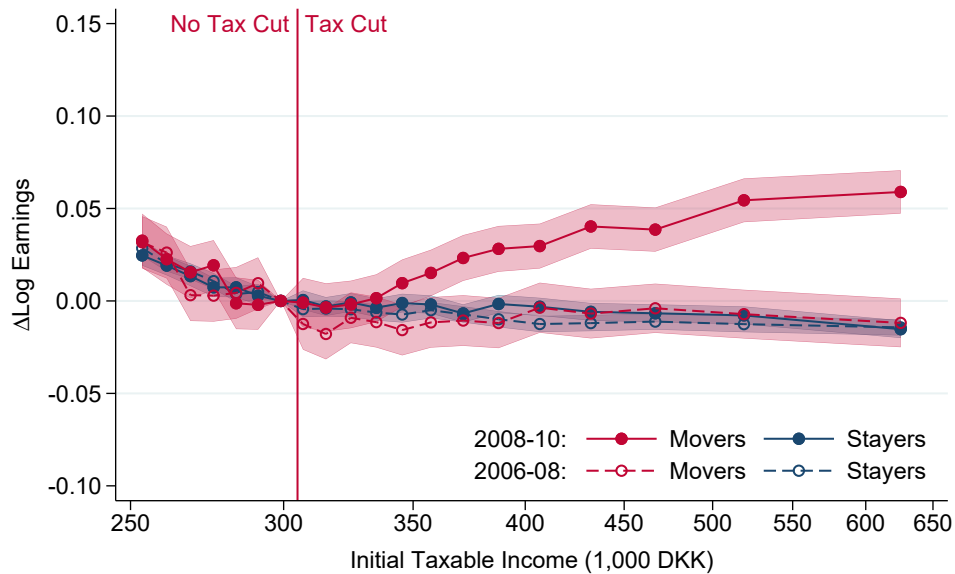


**FIGURE A.5: IMPACT OF TAX REFORM ON EARNINGS**  
BY TYPE OF SWITCH

**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 C), 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 previous figures. It plots changes in log earnings between 2008-10 (post-reform) and between 2006-08 (pre-reform placebo) by income bin, omitting the bin just below the treatment threshold depicted by the vertical line. The empirical patterns are quite similar in the two panels. The earnings responses to lower taxes are clear and sizeable regardless of the type of switch. 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:

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

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

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

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

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

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

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

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

$$\text{corr}(z_t, y_t) = \frac{\text{cov}(z_t, y_t)}{\sigma_{z_t} \sigma_{y_t}}, \quad (21)$$

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

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

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  $\sigma_{z_t}^2 = \lambda \sigma_{y_t}^2 + (1 - \lambda) \sigma_{z_{t-1}}^2$  from the earnings specification (2). This implies that  $\sigma_{z_t}^2 = \lambda \sum_{s=0}^{\infty} (1 - \lambda)^s \sigma_{y_{t-s}}^2$ . From equation (4) and  $n_t = f(t) + \mu$ , it follows that  $\sigma_{y_t}^2$  is time-invariant, i.e.  $\sigma_{y_t}^2 = \sigma_y^2$  for  $\forall t$ . Using this time-invariance along with the property  $\sum_{s=0}^{\infty} x^s = \frac{1}{1-x}$ , it follows that  $\sigma_{z_t}^2 = \sigma_y^2$  and, hence,  $\sigma_{z_t} = \sigma_y$ . By inserting this property and the above formula for the covariance into the definition in (21), we obtain

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

### 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  $\rho$  is the social discount factor. This objective is not well-defined for  $\rho = 1$  and, therefore, we redefine the planner's objective function as

$$\Psi \equiv (1 - \rho) \sum_{t=0}^{\infty} \rho^t D_t. \quad (22)$$

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

$$\begin{aligned} \Psi &= D^* + (1 - \rho) \sum_{t=0}^{\infty} \rho^t (D_t - D^*) \\ &= D^* + (1 - \rho) \sum_{t=0}^{T-1} \rho^t (D_t - D^*) + (1 - \rho) \sum_{t=T}^{\infty} \rho^t (D_t - D^*) \end{aligned} \quad (23)$$

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 (23), 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

$$\begin{aligned} \lim_{\rho \rightarrow 1} \left| (1 - \rho) \sum_{t=0}^{\infty} \rho^t D_t - D^* \right| &\leq \lim_{\rho \rightarrow 1} (1 - \rho) \sum_{t=0}^{T-1} \rho^t (D_t - D^*) + \lim_{\rho \rightarrow 1} |D_T - D^*| \rho^T \quad \forall T \\ \Leftrightarrow \quad \lim_{\rho \rightarrow 1} \left| (1 - \rho) \sum_{t=0}^{\infty} \rho^t D_t - D^* \right| &\leq |D_T - D^*| \quad \forall T. \end{aligned}$$

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

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

Therefore, at a social discount factor of  $\rho = 1$ , the welfare objective in equation (22) 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):

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

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

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

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

## B.5 Endogenous $\lambda$

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  $\lambda$ . Considering a steady state with constant productivity  $n$  and effort  $y$  (to simplify exposition), we solve for the constrained-efficient solution of  $(y, \lambda)$  that maximizes worker-firm surplus.<sup>29</sup> The per-period surplus is given by

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

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  $\lambda$  equals

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

With costless verification ( $q = 0$ ), we have  $v'(y/n) = 1 - \tau$ . Given the parameterization  $v(x) = \frac{\eta}{\eta+1} x^{\frac{\eta+1}{\eta}}$  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  $v'(y/n)$ .

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

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

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

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

---

<sup>29</sup>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  $\lambda$ :

$$\frac{\lambda}{1-\lambda} = \frac{\eta(1-\delta)^2}{\gamma(1-(1-\lambda)\delta)^2}, \quad (24)$$

where  $\gamma \equiv q/y$  denotes the evaluation cost in proportion to output. We may interpret  $\gamma$  as capturing the degree/cost of imperfect information, which determines where  $\lambda$  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  $\gamma$ , the evaluation frequency  $\lambda$  lies between 0 and 1, thereby giving rise to the dynamic return mechanisms characterized in this paper. As for comparative statics, equation (24) shows that  $\lambda$  is decreasing in the evaluation cost  $\gamma$ , increasing in the effort elasticity  $\eta$ , decreasing in the discount factor  $\delta$ , and independent of  $\tau$ . The last result relies on the (natural) assumption that evaluation costs are tax deductible.