Online Appendix:

"Micro vs Macro Labor Supply Elasticities: The Role of Dynamic Returns to Effort"

Henrik Kleven
Claus Kreiner
Kristian Larsen
Jakob Søgaard

March 2025

A Supplementary Tables and Figures

TABLE A.1: TOP VS BOTTOM OCCUPATIONS
JOB TITLES AND EARNINGS

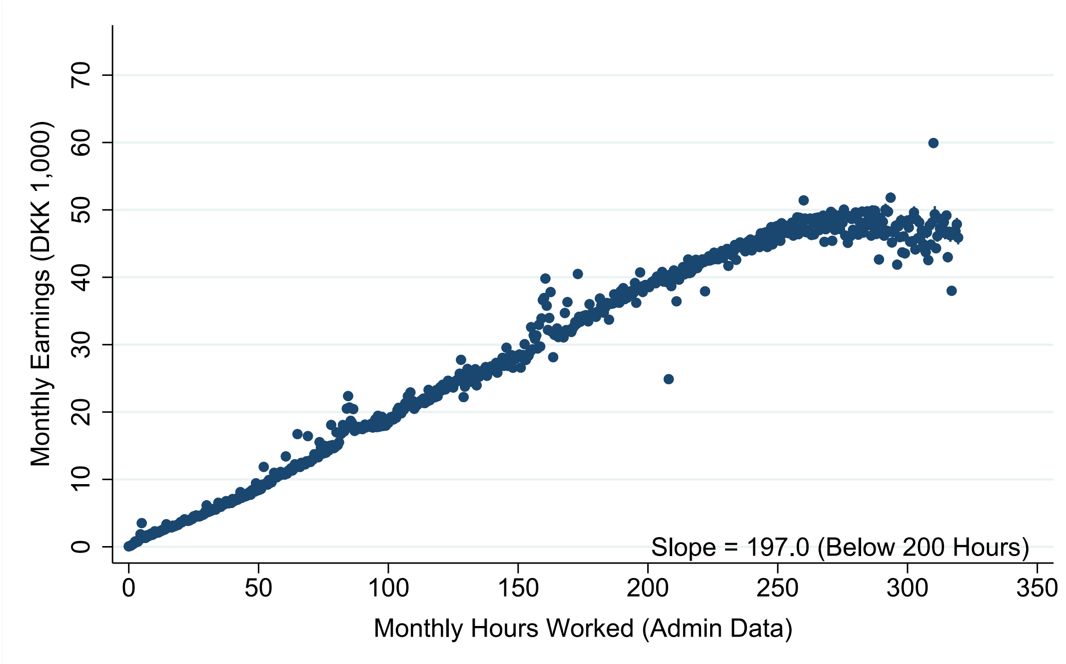
	Occupation	Earnings (DKK 1,000)		
		Mean	**P10**	**P90**
Top 10	Top Executives	2,107.6	447.7	5,181.2
	Managing Directors	1,071.7	407.8	1,890.2
	Securities and Currency Traders	1,050.7	397.5	1,963.2
	Administrative Directors	1,024.0	382.2	1,803.6
	Lawyers	983.6	401.5	1,810.8
	Pilots	929.7	476.7	1,400.5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

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

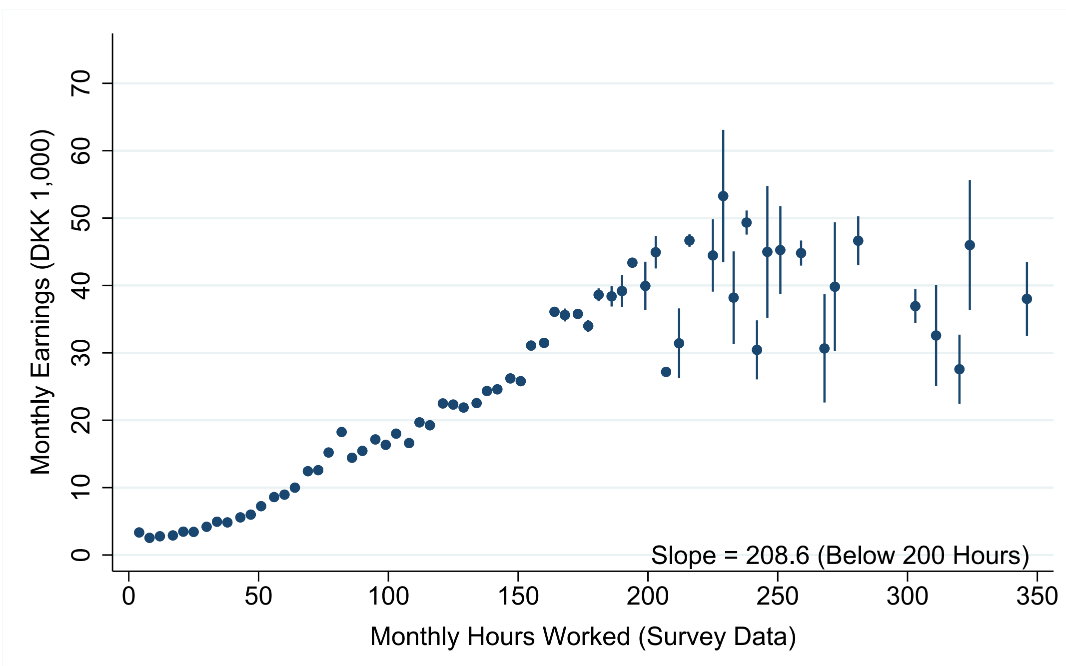
Figure A.1: Validation of Administrative Hours Worked Measure

Administrative Data vs Survey Data

A: Administrative Data (Pension Measure of Hours Worked)

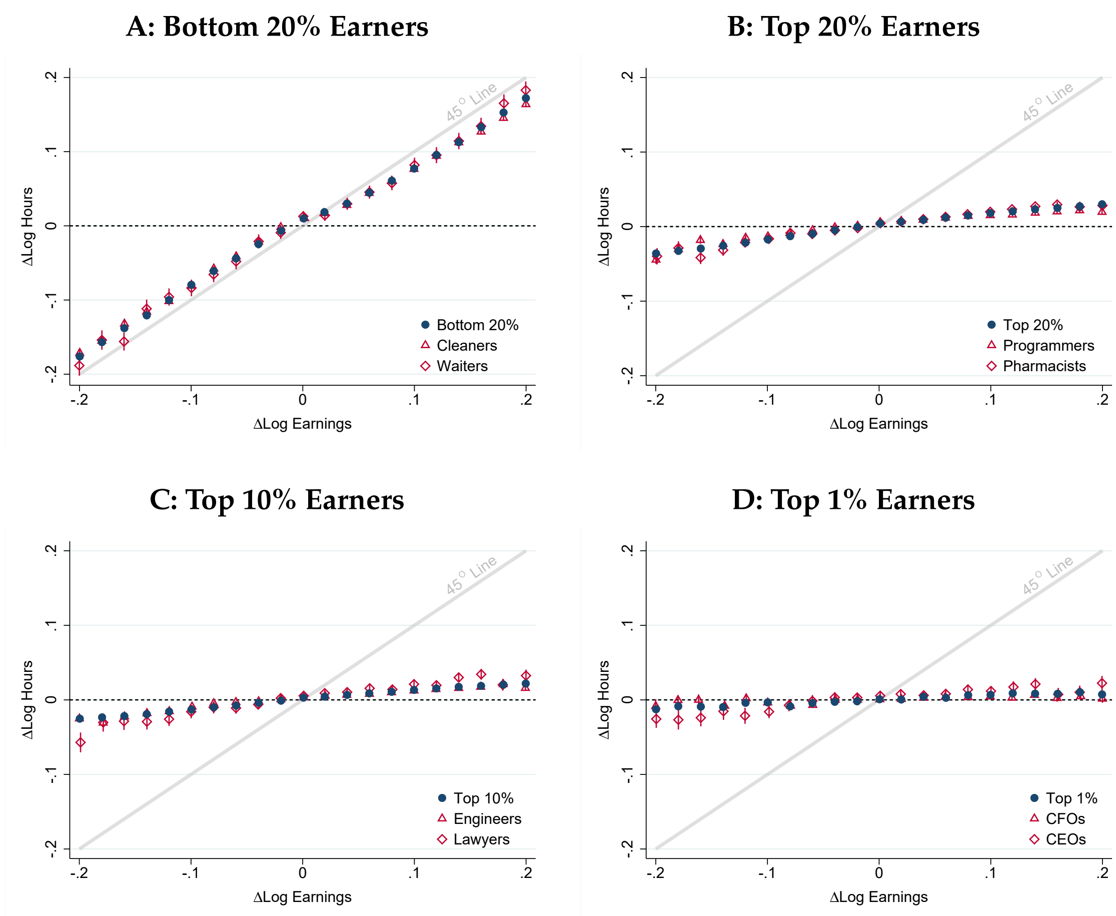


B: Survey Data (Self-Reported Measure of Hours Worked)



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

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

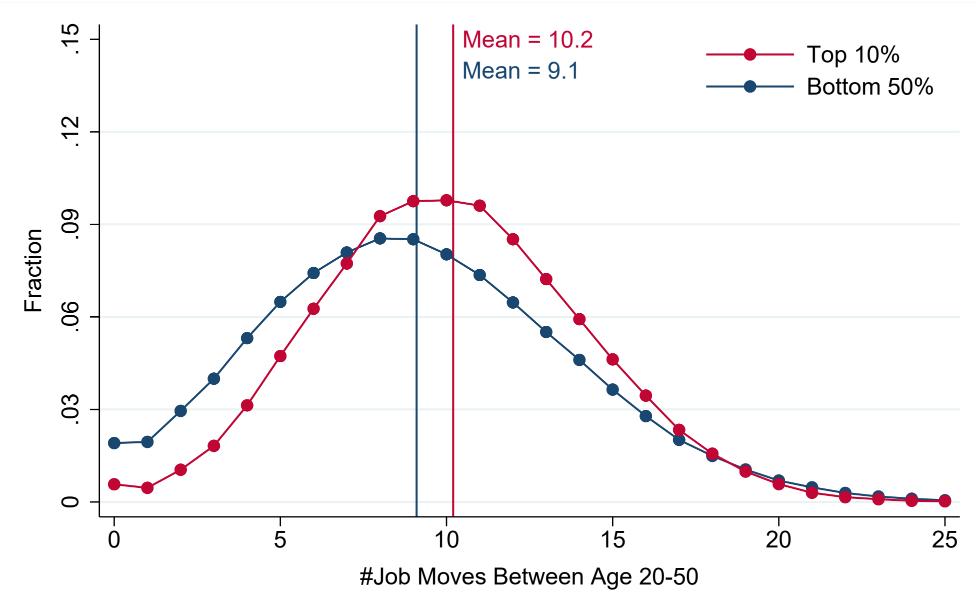


Notes: This figure shows the contemporaneous relationship between hours and earnings changes at the intensive margin. It plots changes in log hours against changes in log earnings in different segments of the earnings distribution: the bottom 20%, the top 20%, the top 10%, and the top 1%. The average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

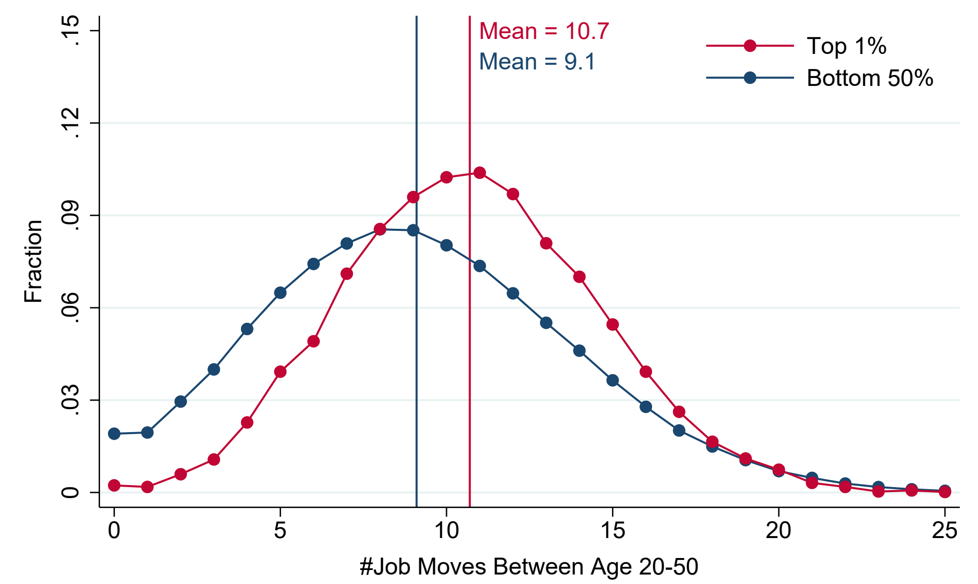
FIGURE A.3: DISTRIBUTION OF NUMBER OF SWITCHES

TOP VS BOTTOM EARNERS BETWEEN AGES 20-50

A: Top 10% vs Bottom 50%



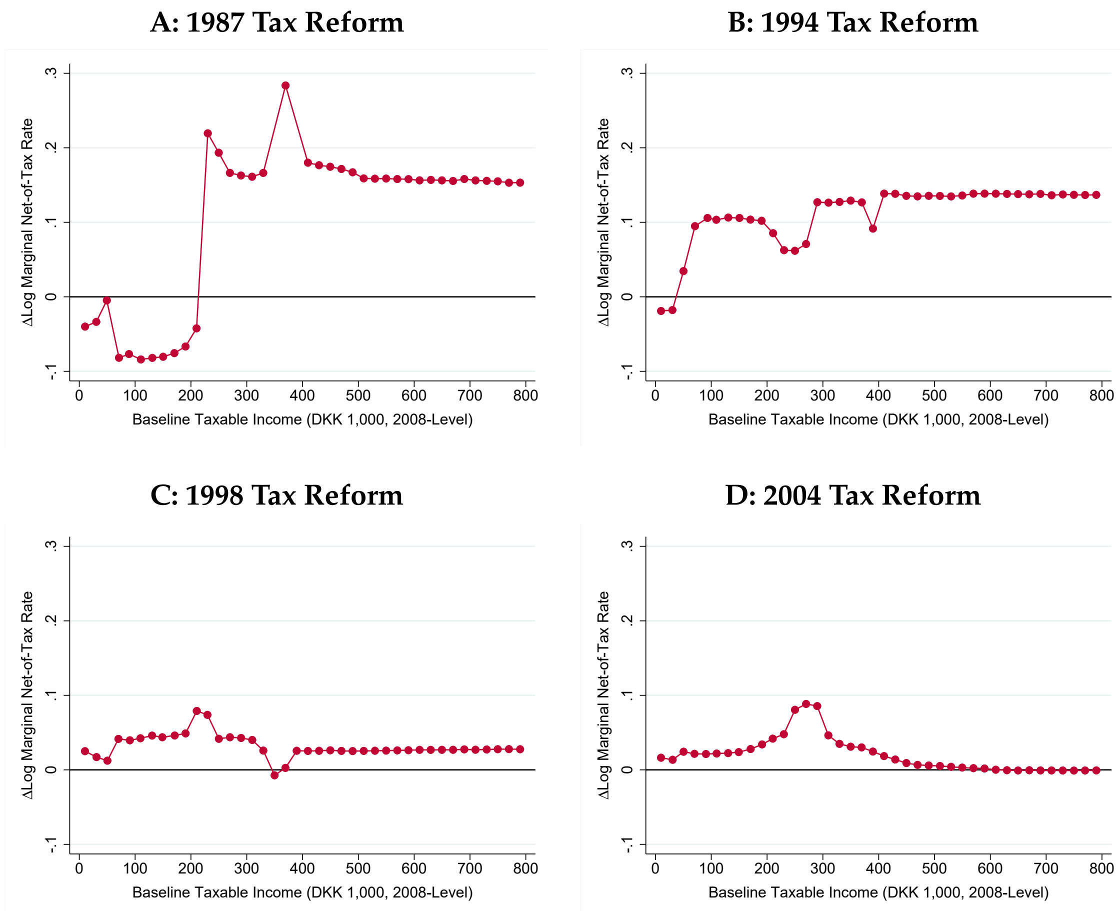
B: Top 1% vs Bottom 50%



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

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

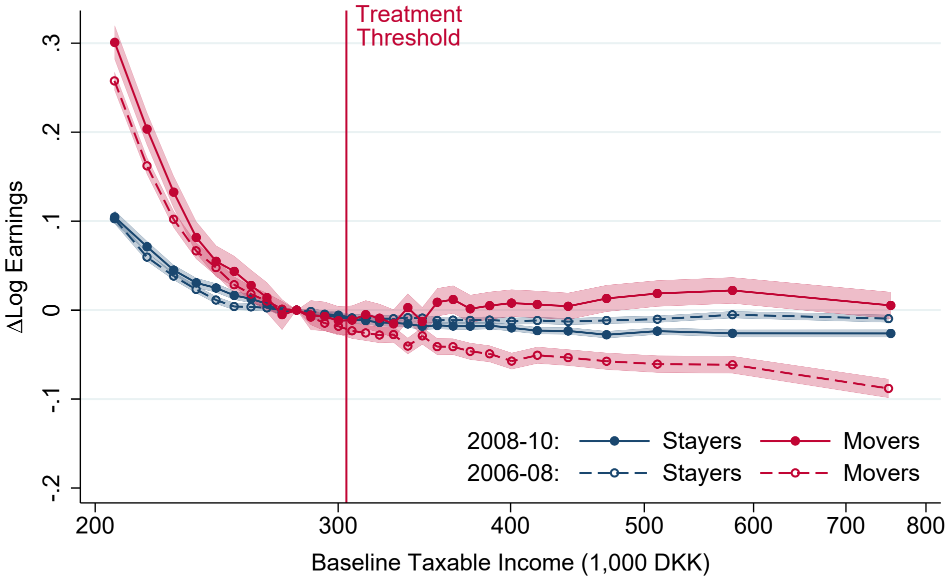
REFORM-INDUCED CHANGES IN THE MARGINAL NET-OF-TAX RATE



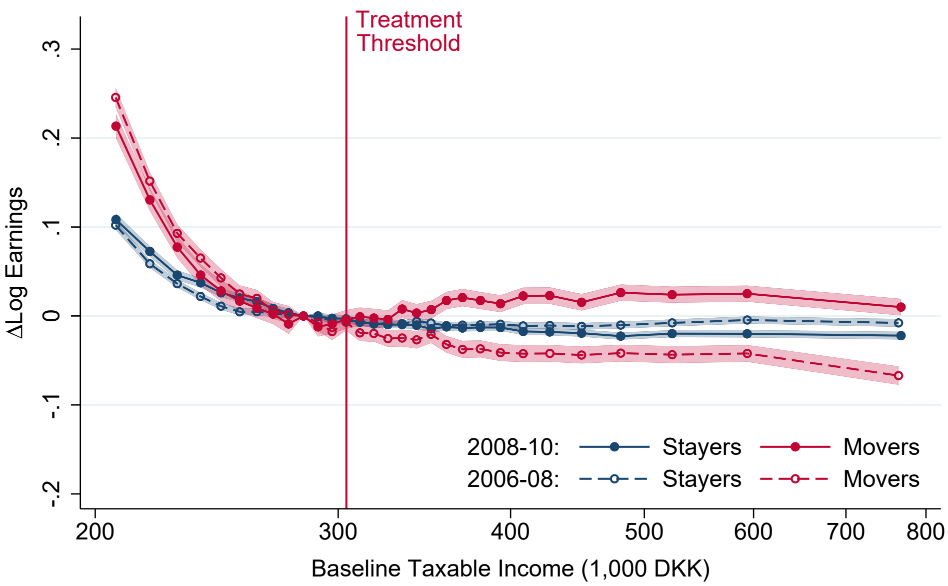
Notes: This figure illustrates the tax variation created by four major tax reforms implemented in Denmark prior to the 2009 reform. Each panel plots the change in the log marginal net-of-tax rate $1 - \tau$ by income bin. The general theme of Danish tax reforms since the 1980s has been to lower marginal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

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

A: Firm Switches



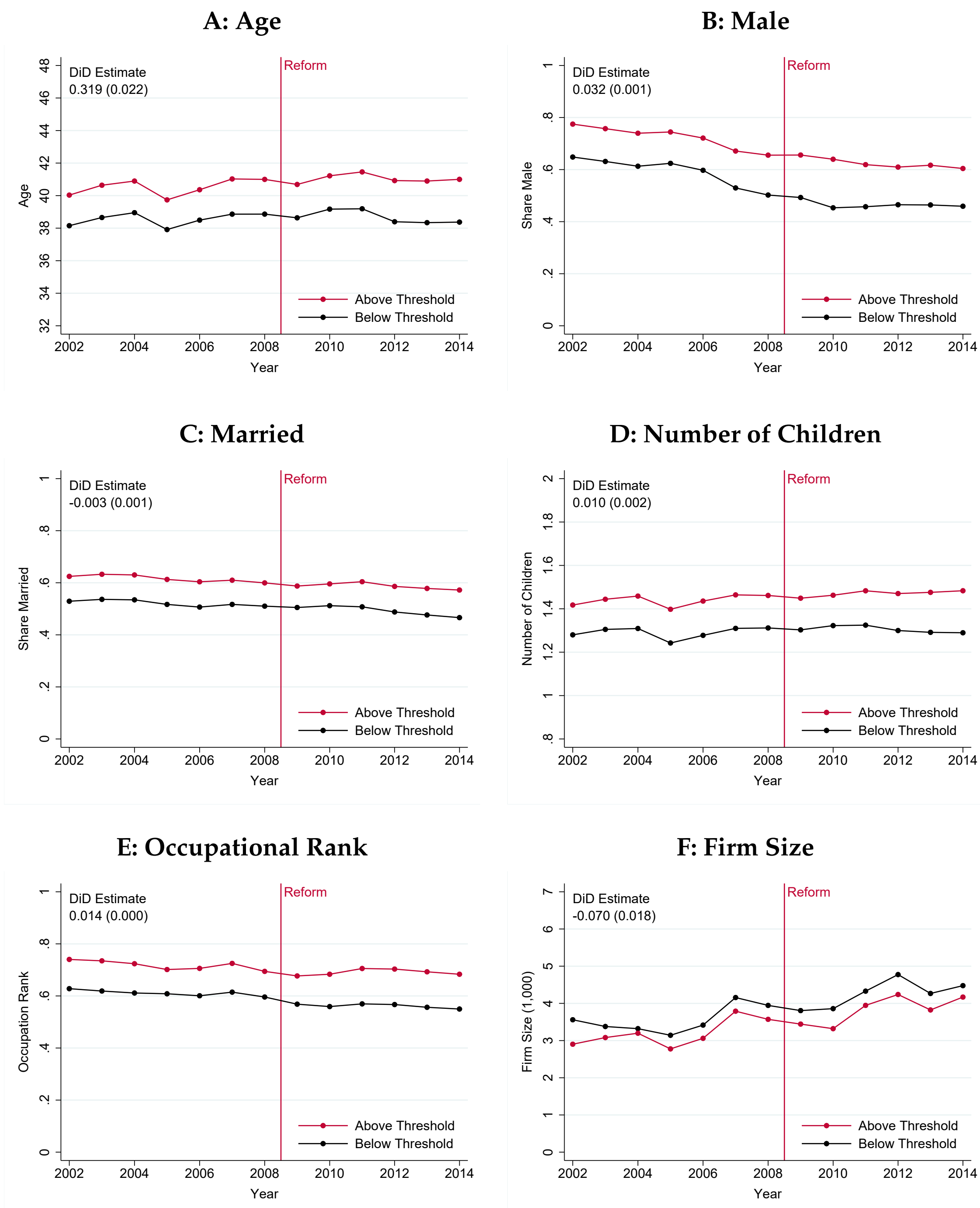
B: Occupation Switches



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

Figure A.6: Is Switching Selected?

Impact of Tax Reform on Switcher Characteristics

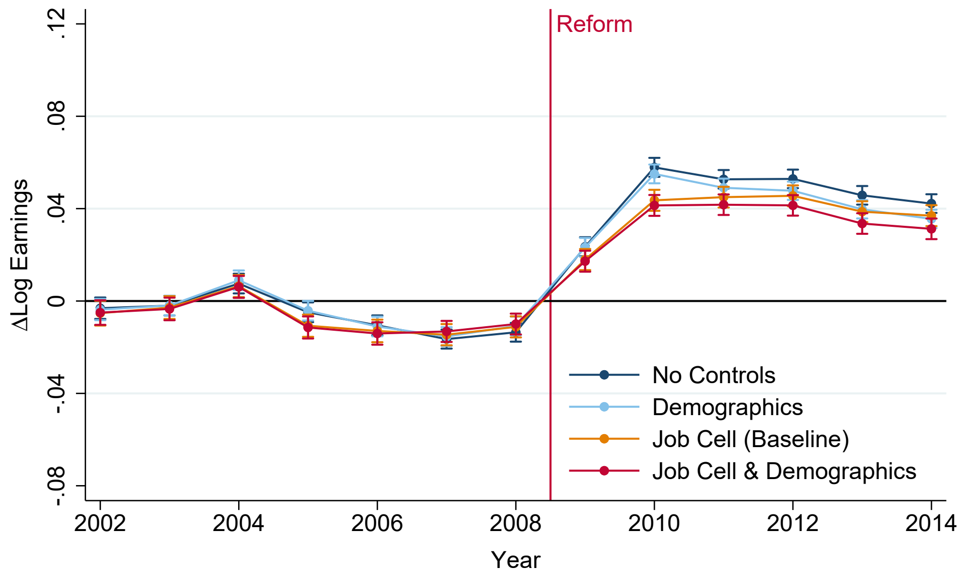


Notes: This figure investigates if the observable characteristics of treated job switchers relative to untreated job switchers diverge after the 2009 tax reform. Each panel plots the time series of a demographic variable for job switchers above and below the treatment threshold. Six different variables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

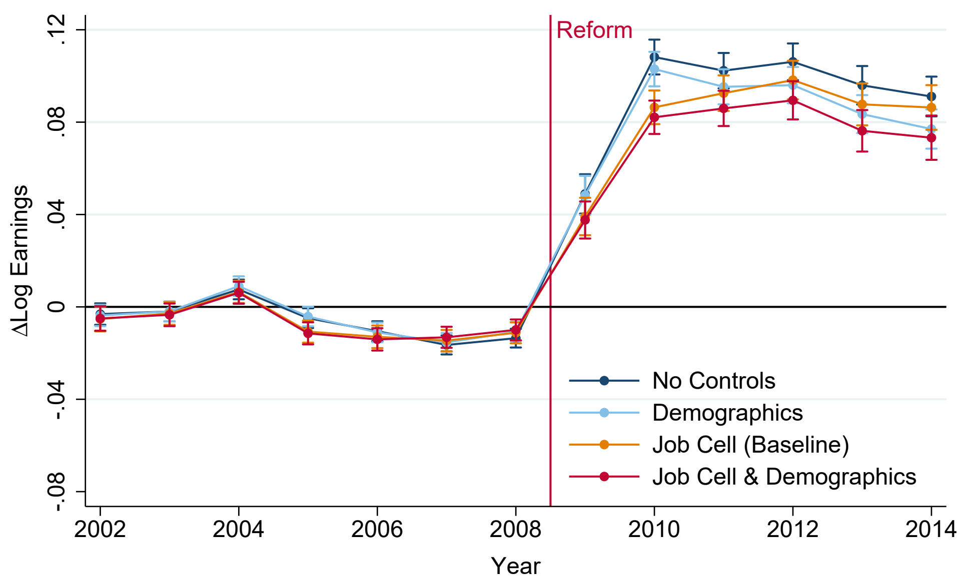
FIGURE A.7: IMPACT OF TAX REFORM ON EARNINGS OVER TIME

ROBUSTNESS OF DYNAMIC APPROACH TO CONTROLS

A: Intention-to-Treat (ITT)

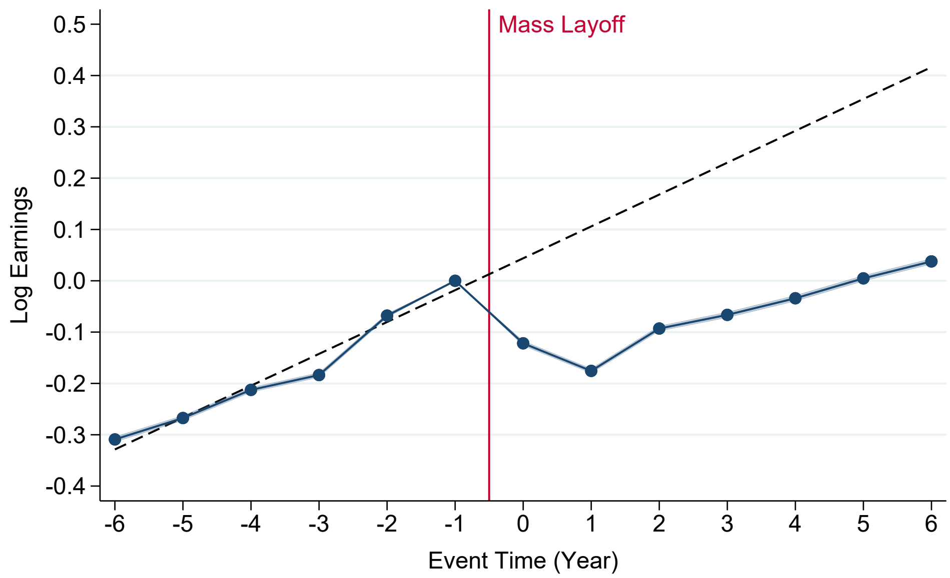


B: Treatment-on-the-Treated (TOT)



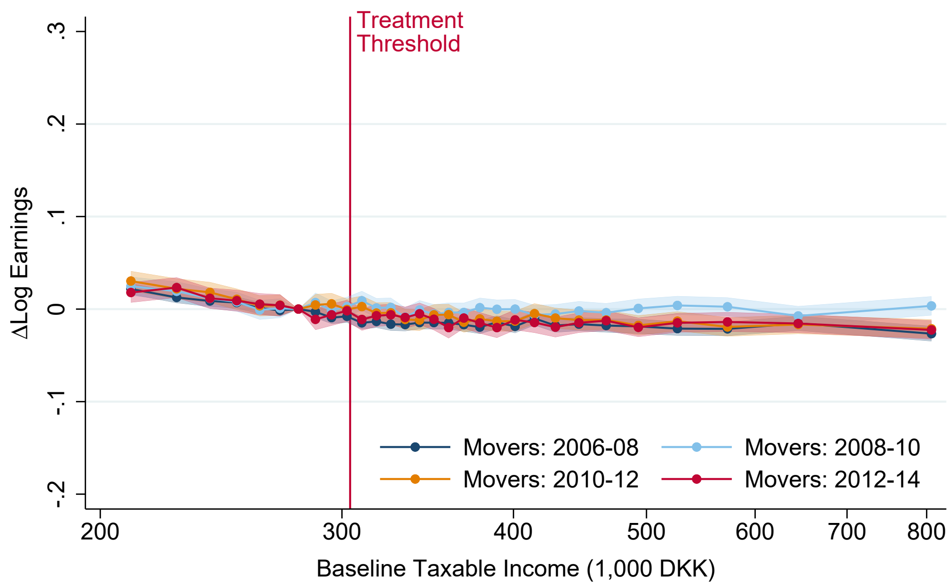
Notes: This figure investigates the robustness of the results in Figure 7. It plots the time profile of earnings effects (ITT and TOT, respectively) among job switchers under different sets of controls. Four different specifications are considered: no controls (dark blue), demographic controls (light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

FIGURE A.8: IMPACT OF MASS LAYOFF ON EARNINGS



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

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



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

B Theoretical Proofs

B.1 Proof of Proposition 1

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

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

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

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

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

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

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

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

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

B.2 Proof of Proposition 2, Part 2

The correlation coefficient between z_t and y_t equals

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

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

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

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

To compute the correlation coefficient, we also use that $\sigma_{z_t}^2 = \lambda \sigma_{y_t}^2 + (1-\lambda)\,\sigma_{z_{t-1}}^2$ 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 $n_t = g(t) + \mu$, 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 (30), we obtain

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

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

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

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

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

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

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

By substituting this into equation (32), we obtain

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

This implies

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

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

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

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

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

B.4 Generalization of Proposition 3

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

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

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

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

where

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

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

B.5 Endogenous λ

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

which gives

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

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

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

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

B.7 Derivation of Equation (18)

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

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

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

B.8 Generalized Earnings Dynamics for Movers and Stayers

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

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

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

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

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

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

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

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

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

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

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