NBER WORKING PAPER SERIES

USING PAYROLL TAX VARIATION TO UNPACK THE BLACK BOX OF FIRM-LEVEL PRODUCTION

Youssef Benzarti
Jarkko Harju

Working Paper 26640
http://www.nber.org/papers/w26640

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020

Harju has received funding from the Strategic Research Council (SRC) at the Academy of Finland No. 293120, Work, Inequality, and Public Policy. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Using Payroll Tax Variation to Unpack the Black Box of Firm-Level Production
Youssef Benzarti and Jarkko Harju
NBER Working Paper No. 26640
January 2020
JEL No. H20,H22,H23

ABSTRACT

This paper uses quasi-experimental variation in payroll taxes to estimate their incidence and investigate how firms use their input factors. We find that higher payroll tax rates lead to large employment responses and have no effects on employee earnings. As payroll taxes increase, firms substitute away from low-skilled, routine and manual workers towards more productive workers and also reduce investments. Our results imply that, contrary to the canonical tax incidence model, firm-level production and input factor choices are affected by payroll taxes.

Youssef Benzarti
Department of Economics
University of California
Santa Barbara, CA 93106
and NBER
youssef.benzarti@gmail.com

Jarkko Harju
VATT Institute for Economic Research
Arkadiankatu 7
P.O.Box 1279
00101, Helsinki
Finland
jarkko.harju@vatt.fi

1 Introduction

Payroll taxes are ubiquitous: the vast majority of OECD countries and most of the rest of the world impose some form of tax on payroll shared between employers and employees and used to fund social insurance programs. Payroll tax rates can also be substantial, reaching as much as 60 percent in France, for example. Moreover, the share of total tax revenue raised by payroll taxes has been steadily increasing in OECD countries since the 1960's, so much so that payroll taxes now raise more revenue than income taxes (see Figure 1).[1] For these reasons, payroll taxes can impose a substantial burden on the economy. The consensus among economists is that this burden is mostly borne by workers, and therefore the price of labor faced by firms is mostly undistorted by payroll taxes.

In this paper, we estimate the incidence of payroll taxes and, in doing so, open the black box of the firm to assess how much payroll taxes bias the use of production factors. We use unique variation in the employer portion of payroll tax rates in Finland. Finnish employers face a discontinuous increase in payroll tax rates if they exceed a set depreciation threshold. This triggers a significant change in payroll tax rates, which is equivalent, on average, to a 5 percentage point increase in corporate taxes. Importantly, this variation affects all employees in the firm, irrespective of their age, occupation status, etc., but does not affect the benefits they are entitled to. This distinguishes our paper from recent payroll tax incidence papers such as Saez et al. [2012], Saez et al. [2019] and Bozio et al. [2019] and ends up playing an important role in interpreting our results.

Using this exogenous variation, we first establish that payroll taxes do not affect net-of-payroll-tax employee earnings, implying that firms bear the burden of payroll taxes. We then estimate the causal effect of payroll taxes on employment and find that payroll taxes reduce the number of employees, but with substantial heterogeneity by skill level and type of occupation: the employment effects of payroll taxes are concentrated among low-skilled

[1] In the US, for example, the share of Federal revenue raised by payroll taxes has increased from less than 10 percent in the 1950's to more than 34 percent in 2016. In OECD countries, they raise 26 percent of total tax revenue, which is higher than the revenue raised by personal income taxes.

workers and workers performing routine tasks. We also estimate that payroll taxes tend to reduce investments, which could be either consistent with workers and capital being complements or with scale effects. Our empirical results support the existence of some scale effects which tend to mask the substitution between labor and capital.

There are two main potential concerns with our empirical approach, which we address. First, regression discontinuity designs can be sensitive to functional form assumptions as well as bandwidth choice. We address these issues in several ways. First, we use a placebo test that relies on years when the discontinuity in payroll taxes did not exist and estimate small and statistically insignificant discontinuities in outcomes at the threshold. Second, we provide plots of data around the cutoff that transparently and non-parametrically show the presence of a discontinuity in our outcomes of interest. Third, we show that our estimates are not sensitive to the use of different functional forms. Fourth, we use the optimal bandwidth estimates from Calonico et al. [2014], but also vary the bandwidth and find that our estimates are not affected by bandwidth choice.

The second potential concern is that we might be estimating evasion responses rather than real responses. While firms bunch at the threshold, which could be due to evasion, we do not use this bunching in our estimation. In addition, firms could be misreporting their number of employees, by hiring them under the table, to avoid the additional payroll tax. There are three reasons why we believe this behavior cannot explain our results. First, our dataset contains both accounting and tax measures of the variables we observe and, the accounting variables are systematically third-party-audited: 93 percent of firms in our baseline sample are subject to such audits. If our results were due to evasion, our estimates should be substantially smaller when using the accounting variables instead of the tax variables and yet we estimate very similar effects in both cases. Second, if the employment response we estimate was due to firms evading payroll taxes, we should not expect investments to negatively respond to the increase in payroll taxes as well, but we observe clear effects on these outcomes. Third, under-reporting the number of employees is very unlikely in

Finland: while firms would be saving on payroll taxes, black market employment would deprive employees of all social insurance benefits and would put them and their employers at risk of facing legal consequences. Reducing wages to compensate for an increase in payroll taxes would be far less risky/costly. For these three reasons, it is very unlikely that our estimates are due to evasion and instead we are likely estimating real responses.

This paper contributes to the following literatures. The first literature we contribute to is the tax incidence literature. First, since taxes are at the heart of redistribution, knowing whether they affect different skill levels differently is key for the design of optimal tax systems. We provide some of the first evidence on this question by showing that payroll taxes affect workers of different skill levels employed in different tasks differently. Second, we show that payroll taxes tend to depress investment either through a capital-labor complementarity or because of liquidity and scale effects, which should be accounted for when scoring payroll tax changes. The US Congressional Budget Office, for example, currently assumes, as is standard in the tax incidence literature, that payroll taxes are fully borne by workers.[2] This assumption implicitly implies that payroll taxes do not distort firm-level input use, which is inconsistent with our findings and highlights the potential policy implications of our paper. Third, we complement the compelling evidence of Saez et al. [2012] and Saez et al. [2019], who use employee age- and cohort-based variation in payroll tax rates to question the consensus that payroll taxes are borne by workers. Given that their identifying variation is age- and cohort-specific, they argue that pay inequality concerns could explain their finding that payroll taxes are borne by firms, since otherwise employers would be paying two different wages to workers of different ages/cohorts but who are otherwise similar. In contrast, we first show that, even in settings where payroll tax changes apply to all workers in a given firm, which circumvents any issues of pay inequality between two workers within the same firm, payroll taxes are still borne by firms, further exacerbating the inconsistency of this finding with the canonical tax incidence model. Second, we can assess the distributional effects of

[2]See, for example: https://www.cbo.gov/budget-options/2018/54805

payroll taxes across the skill and task spectrum, which is not easily implementable in Saez et al. [2012] and Saez et al. [2019] because of the nature of their identifying variation. Third, our identifying variation allows us to disentangle two possible channels through which payroll taxes affect firm-level outcomes: (1) the liquidity channel, i.e. firms changing their behavior because of the liquidity constraints that higher payroll taxes impose on them, and (2) the marginal cost channel, i.e. firms changing their behavior because payroll taxes distort the marginal cost of labor. This is also difficult to implement in Saez et al. [2012] and Saez et al. [2019] because of the nature of their variation. Saez et al. [2019], for example, compare firms with a high share of young workers versus firms with a medium share of young workers, which, as they acknowledge, would capture both the liquidity and marginal cost channel effects at the same time. In our paper, we find evidence more consistent with the liquidity channel.

Second, while there is a large body of work discussing job polarization and its effects and causes (see Autor et al. [2006] and Goos et al. [2009]), there is limited evidence on how taxes affect the relative distribution of workers across the skill spectrum. Our paper is one of the first to show that payroll taxes affect skill levels and job tasks very differently. We believe this is important, both because we provide an additional channel that could affect job polarization that had not been explored before and also because our findings show that payroll taxes, possibly differentiated by skill group or by task, could be used as a policy tool to counteract job polarization.

Finally, we contribute to the literature that estimates the capital-labor elasticity of substitution. The debate in this literature has mostly centered around whether the capital-labor elasticity of substitution is greater than, equal to or smaller than 1 when using a constant elasticity of substitution (CES) production function. The consensus has been that the elasticity is equal to 1, prompting researchers to use a Cobb-Douglas production function. More recently, this consensus has been questioned, for example by Raval [2014] and Oberfield and Raval [2014], who estimate a capital-labor elasticity that is smaller than 1 (but larger than zero), using a CES framework. Since we estimate that both capital and labor decrease when payroll taxes increase, our evidence is consistent with a capital-labor elasticity, at the micro level, that is equal to zero, i.e., capital and labor are estimated to be complements in the CES framework. However, this decrease in both capital and labor could also be consistent with liquidity effects, which are implicitly assumed away in the CES framework. For this reason, *at the micro level*, our findings could be either consistent with capital and labor being complements or with liquidity effects dominating the capital labor substitution effect. Using an empirical test, we show that the liquidity effects are likely to dominate, which calls for the literature to incorporate and investigate them.

2 Institutional Background and Data

2.1 Institutional Background

2.1.1 Payroll Taxes in Finland

In Finland, social insurance contributions are used to fund pensions, unemployment insurance, accident insurance, health insurance and life insurance. Both employees and employers contribute to social insurance.[3] In general, the largest share of total social insurance contributions goes to pension contributions and employers' statutory share of total contributions is larger than that of their employees. For example, in 2017, the average pension insurance contribution rate was 17.95 percent of a given employee's monthly gross wage and the employee's contribution rate was 6.15 percent. In this paper, we use variation in how much employers have to contribute to their employees' health and pension fund.

[3] The split between employees and employers depends on several firm and worker characteristics, including, for example, the age of the worker.

2.1.2 Identifying Variation in Payroll Tax Rates

Prior to 2010, there were three employer payroll tax rate brackets for health and pension contributions, depending on the level of capital depreciation and labor costs of the firm, as shown in Table 1 below.[4] Importantly, the contribution rates of employees and the benefits they qualify for were unaffected by these discontinuities. Category I corresponds to firms with less than 50,500 euros of annual capital depreciation or more than 50,500 euros but less than 10 percent of annual salaries. Category II corresponds to firms with depreciation levels of more than 50,500 euros and 10 to 30 percent of labor costs. When depreciation levels exceed 50,500 euros *and* 30 percent of labor costs, contributions are paid according to Category III.[5] The rationale for these three categories was to support labor-intensive firms by reducing their labor costs and they were originally introduced in April 1973.

Employers' payroll tax rates are an increasing step function of the category that firms belong to. We focus on comparing firms in Category I, which fall below the depreciation threshold, to firms in Categories II and III, which are above the threshold. As illustrated in Appendix Table 14, firms in Categories II and III face a systematically high payroll tax rate, approximately 2 to 3 percentage points higher, depending on the years and the category.

Table 1: Firm categories for payroll tax rates

Definition for firm categories					
I	D < 50,500	**or**	D ≥ 50,500	and	D < 0.1 * labor costs
II	D ≥ 50,500	and	D ≥ 0.1 * labor costs	and	D < 0.3 * labor costs
III	D ≥ 50,500	and	D ≥ 0.3 * labor costs		

Note: D refers to tax-deductible capital depreciations and labor costs refer to all salaries.

In January 2010, these three categories were abolished and the three different contribution rates were replaced with one single rate for all firms irrespective of annual capital depreciation levels and labor costs.[6]

[4] We provide details of the depreciation rules in Appendix Section A.

[5] These categories were determined by the latest available tax information and salaries paid for the same year as that used to determine the depreciation levels. For example, the 2006 payment category was based on fiscal year 2004.

[6] See legislation in Finlex: Government Proposal 147/2009.

2.1.3 Collective Wage Bargaining

In Finland, minimum wages are negotiated first at the national level and then at the industry level between industry-specific employee unions and industry-specific trade unions. These negotiations set a minimum wage level, which is industry-specific. Importantly, these agreements apply to all workers, not only to employees who belong to labor unions. Although the collective bargaining is extensive (over 90 percent of employees are covered by a collective wage agreement), wages can vary across firms and across employees within firms. Firms can, of course, pay higher wages but also lower wages as long as they remain above the industry-specific minimum wage. Therefore, collective wage agreement could affect the incidence of payroll taxes on wages and earnings, but only for workers for whom the minimum wage is binding, which we assess below.

2.1.4 Accounting versus Tax Depreciation

In Finland, depreciation for tax purposes can differ from depreciation for accounting purposes. Depreciation for accounting purposes is a systematic reduction of the cost of a fixed asset and is subject to strict auditing and is thus difficult for the firm to manipulate. According to Finnish tax law, the amount of annual depreciation for tax purposes cannot be larger than that for accounting purposes. This opens up the possibility for firms to manipulate the amount of tax depreciation, e.g. by reducing tax depreciation to qualify, for example, for a lower payroll tax rate, whereas this type of manipulation is virtually impossible for accounting depreciation. Fortunately, we have data on both the accounting and tax depreciation levels and thus we can examine the extent to which this manipulation exists. In Appendix Figure 14, we show that (1) the accounting and taxation depreciation levels are very highly correlated (upper panel of Figure 14), and (2) the excess mass of bunching firms at the threshold is very similar when using the accounting and taxation variables (lower panel of Figure 14). We also implement our empirical approach, outlined below in Section 3, on both the accounting and tax measures and find very similar results.[7]

2.2 Data

We use firm-level tax record data covering the universe of Finnish firms from 1996 to 2015, provided by the Finnish Tax Administration. The dataset contains a rich set of firm-level variables and firm characteristics, including organizational form, location and industry code. The dataset provides yearly information, at the firm level, on labor costs, number of employees, both accounting and tax amounts of capital depreciation and the level of capital investment. Importantly, we can separate investments into three different main categories: fixed assets, buildings and research and development. In addition, we have firm-level data on sales and various cost categories, including material and rental costs.

The upper panel of Appendix Table 15 shows summary statistics for the main firm-level variables used in the empirical analysis. The lower panel of Table 15 shows the same summary statistics for the whole population of Finnish firms with annual sales ranging between 10,000 and 100,000,000 euros, to illustrate how comparable the sample of affected firms is to the whole population of firms. On average, firms are larger in our sample in comparison to all Finnish firms, but the difference is relatively small. In Appendix Table 16, we also show the distribution of industry and organizational forms. These two distributions are well balanced across the threshold. Overall, the variation we use relies on a sample of firms that is reasonably representative of the population of firms in Finland.

The only data restriction we apply throughout the paper is that we exclude all firms that were not subject to the depreciation rules we consider. Specifically, we remove all firms that have capital depreciation below 10 percent of all wages. Legally, the discontinuity in payroll tax rates we consider does not apply to these firms, so there is no reason to include them in the analysis. This restriction removes approximately 25 percent of the total data.

[7]The results using the tax measures, which are our baseline results, are discussed in Section 3 and the results using the accounting measure can be found in Appendix Table 13 and are displayed in Appendix Figure 15.

In addition to the firm-level data, we also observe a wide set of information on the employees of these firms, and, importantly, we can match the employee-level dataset to the firm-level dataset using a unique identifier. Employee-level data are reported annually, are based on job contracts and contain the following variables, among others: gender, age, working days and months and annual earnings in each firm. These data also include the starting and ending dates for each employee-firm job contract pair. In addition, we link the employee-level dataset to an administrative dataset that contains information on the education levels of all Finnish individuals in two forms: (1) a dummy for whether an employee has a high school or a vocational school diploma, and (2) a six-category classification of the highest education level attained. We first link the job contract data (which contain information on employees and unique firm identifiers) to our firm-level tax register data containing the annual depreciation levels and other yearly firm-level variables. We are able to match 93.2 percent of all firm-year pairs to their employee-year pairs. Second, we link these two datasets to the dataset containing employee education variables, with a match rate of 99 percent. Appendix Table 17 provides definitions of the outcome variables and other variables used in the empirical analysis below.

3 Empirical Strategy

To estimate the response of capital investment to labor costs, we use a discontinuity in payroll tax rates at the €50,500 depreciation threshold as described in Section 2.1. As firms cross the €50,500 depreciation threshold, the average (and marginal) payroll tax rates discontinuously increase, effectively increasing labor costs, as shown in Figure 2.

Our empirical analysis proceeds in four steps. First, we provide graphical evidence by plotting all our outcomes of interest around the payroll tax discontinuity to ensure that any estimated discontinuity in these outcomes is graphically present.

Second, we formally estimate the size of the discontinuity in our outcomes of interest around the payroll tax rate threshold using a donut hole regression discontinuity design. Because our running variable (depreciation levels) can be manipulated by firms, we exclude firms that bunch at the cutoff. The upper panel of Figure 3 shows that firms adjust their depreciations to avoid exceeding the threshold. Therefore, we cannot use a standard regression discontinuity design (RDD) approach to estimate the response of capital investment to labor. Instead, we use a donut hole regression discontinuity design, as in Bajari et al. [2011], Card and Giuliano [2014] and Barreca et al. [2016]. We use the method from Kleven and Waseem [2013] to determine the manipulated area which, in their framework, corresponds to the area of the excess and missing masses. We describe this approach in detail in Appendix Section B.

After defining the donut hole region using the bunching method, we follow the approach of Calonico et al. [2014] to estimate the mean square error optimal bandwidth and report bias-corrected estimates with robust standard errors. In addition, we perform placebo tests by running our specification on the post-2010 years, after the repeal of the payroll tax discontinuity. Formally, we run the following regression:

$$\log(y_i) = \alpha + \beta_1 \cdot (depr_i - d) + \beta_2 * Above_i + \beta_3 * Above_i * (depr_i - d) + \epsilon_i \tag{1}$$

where y_{it} is the outcome of interest for firm i, $depr$ is the level of capital depreciations, d is the depreciation threshold above which the average payroll tax rate increases, $Above$ is a dummy (1 above the depreciation threshold, 0 otherwise), ϵ_i is the error term, which is estimated following Calonico et al. [2014] and β_3 is the coefficient of interest showing the magnitude of the change of the outcome variable at the payroll tax rate discontinuity. All variables used in the analysis are defined in Appendix Table 17.

Third, to ensure that our estimates are not spurious, we run equation (1) on the pre-2010 years and the post-2010 years separately for each outcome. The treatment years are the pre-2010 years, when the payroll tax discontinuity was in place. The post-2010 period corresponds to the placebo years, when there was no payroll tax discontinuity. As a result, the post-2010 period offers a plausible falsification test.

Fourth, we perform several robustness checks, including varying the size of the donut hole, the bandwidth and using different degrees of polynomial fit. Our results are robust to all of these checks, which we describe in detail in Section 4.

Note that, in principle, we could use a difference-in-difference (DD) approach to estimate the effect of the repeal of the threshold in 2010. The main reason we do not use this approach is because we lack a clear-cut counterfactual control group. First, when the payroll tax rate threshold was repealed in 2010, the payroll tax rates decreased for all firms, including those below the threshold, which would bias our DD estimates if firms below the threshold respond differently to those above the threshold due to underlying firm heterogeneity in the responses. Second, firms periodically move across the threshold, which poses challenges in defining treatment and control groups and is likely to bias the intensity of treatment in ways that are hard to account for empirically. These issues are not present in our RD donut hole setting, since we only rely on cross-sectional variation.

4 Results

In this section, we first establish that there is indeed a discontinuity in the average payroll tax due by firms. We then use this discontinuity to estimate the effect of payroll taxes on earnings and then on employment. Finally, we consider the effect of payroll taxes on capital, as well as on other firm-level outcomes (including sales and productivity measures).

Payroll Tax Rate. Figure 2 plots the average payroll tax rate for health and pension contributions above and below the €50,500 depreciation cutoff. The average payroll tax rate exhibits a clear discontinuity at the cutoff, with an increase of 2.6 percentage points. This confirms the presence of a discontinuity in payroll taxes and validates our empirical design. While seemingly small in magnitude, especially compared to the payroll tax change analyzed in Saez et al. [2019], this variation is substantial because, contrary to other payroll tax incidence papers, it affects *all employees* in a given treated firm. On average, it corresponds to a 5 percentage point change in corporate taxes for the firms close to the payroll tax rate threshold in our data.

Earnings. Figure 4a plots the effect of the payroll tax discontinuity on individual employee earnings net of the employer and employee portions of payroll taxes.[8] There is no evidence of a discontinuity in earnings at the threshold, implying that employees above the cutoff do not appear to bear the higher payroll taxes. Using equation (1), we estimate the discontinuity in earnings at the threshold both in the treatment sample (years 1996 to 2009) and the placebo sample (years 2010 to 2015). Table 2 shows the corresponding results: we estimate a small and insignificant response in earnings in both the treatment and placebo samples of -0.003 and -0.032, respectively. Similarly, there is no response in earnings to payroll taxes when estimating equation (1) on different earnings deciles, as shown in Appendix Table 9, or on different types of workers (unionization status, gender, education and type of task), as shown in Appendix Table 8. This mitigates concerns that earnings are mechanically prevented from responding because of collective bargaining agreements, which only bind for low-earners, and is a test implemented, for example, in Saez et al. [2019]. We return to this in Section 5.

Labor Costs. Next, we consider the effect of the payroll tax rate discontinuity on labor costs at the firm level. We define labor costs as the total amount spent by a firm on their employees net of the employer and employee portion of payroll taxes. Figure 4b plots the response of labor costs to the discontinuity. We observe a decrease in labor costs just above the cutoff, implying that net of payroll tax labor costs decrease as payroll taxes increase. This is confirmed by the regression estimates, which show a 17.5 percent reduction in labor costs. The corresponding placebo estimate is 7.6 percent.

Since we have estimated that earnings do not respond to the payroll tax, but labor costs

[8] Note that we winsorize the data by dropping the top and bottom 5% of observations in order to remove outliers. Winsorizing does not affect the magnitude of the estimates much but reduces their variance.

do, and since labor costs are roughly the product of earnings and the number of employees, this implies that employment likely responds to the payroll tax discontinuity. Therefore, next, we estimate the effect of the payroll tax discontinuity on employment.

Employment. Figure 4c confirms that the labor costs response is mostly due to a decrease in employment (rather than a decrease in earnings): as payroll taxes increase, the number of employees at a given firm decreases. We estimate a -8.9 percent response to the payroll tax, as show in column (3) of Table 2. This estimate implies labor demand elasticities that vary between -2.90 and 4.16, depending on the years and the firm category we consider, which matter because payroll tax rates vary slightly across years and firm categories. These estimates are consistent with labor demand elasticities estimated in the labor economics literature.[9]

This finding is important and contrasts with the traditional view of the real effect of payroll taxes on wages and employment. Since the common wisdom is that payroll taxes do not affect the price of labor faced by firms as they are passed through to wages, and because labor demand is more elastic than labor supply, we usually do not expect payroll taxes to have employment effects. The fact that payroll taxes distort employment is consistent with the findings of Saez et al. [2019], who show that, when payroll taxes are reduced for workers aged under than 25, firms tend to employ more of them. We complement their compelling findings in two ways: (1) we establish that these employment effects exist even when across-the-board payroll tax changes are implemented, mitigating concerns that these employment effects may be due to the pay inequality concerns of paying a 25-year-old a different wage than a 26-year-old, and (2) we can assess the distributional effects of payroll taxes across the skill and task spectrum, which is not implementable in Saez et al. [2019], since most 25-year-olds hold entry-level jobs that require limited skill levels and experience.

These estimated employment effects indeed mask important dimensions of heterogeneity

[9] See, for example, Lichter et al. [2015] for a survey of labor demand elasticity estimates and Ku et al. [2020] for a recent estimate using payroll tax variation.

along skill levels but also along the type of tasks workers engage in. We describe both below in detail.

Employment effects along the skill dimension. While the labor economics literature has devoted substantial attention to the importance of skills in the labor market, our knowledge of the differential effects of taxes by skill level is still limited in public finance.[10] In order to investigate this response, we break down our sample of workers into high- versus low-skilled. The skill breakdown is based on educational attainment, as is commonly done in the labor economics literature. In the Finnish education system, there are two main levels of academic achievement: graduating from high school and graduating from college. We perform our classification using these two metrics. Our first breakdown classifies workers without a high school degree as low-skilled, and those with a high school degree as high-skilled. The second classification draws the skill division at graduating from college. Figure 5 shows the employment effects of payroll taxes for these four groups. We detect no employment effects for high-skilled workers, whether defined by college or high school graduation, as shown in Figure 5 panels c and d. Instead, all the effects seem to be concentrated among low-skilled workers, as shown in Figure 5 panels a and b.

The graphical evidence is confirmed by our regression estimates in Table 3: the employment response for low-skilled workers is -22.1 percent (no high school degree) and -16.9 percent (no college degree). In contrast, the effects for high-skilled workers are economically small and statistically insignificant. The placebo tests (years from 2010 to 2015) show no response for either low-skilled or high-skilled workers.

Employment effects along the task dimension. A more recent literature has been arguing that the low-skilled/high-skilled categorization masks important heterogeneity and a better suited categorization is one centered around job tasks, as surveyed in Acemoglu and

[10] See the following for examples of the labor market importance of skills: Card and Lemieux [2001], Carneiro and Lee [2011], Goldin and Katz [2007], Katz and Murphy [1992], Goldin and Katz [1998] and Krusell et al. [2000]

Autor [2011].[11] Following this literature and using our dataset, including job descriptions, we categorize workers into three groups: (1) upper-level employees, which include senior officials and upper management, senior officials and employees in research and planning, senior officials and employees in education and training and other senior officials and employees; (2) lower-level employees, including supervisors, clerical and sales workers and other lower-level employees; and (3) routine and manual workers, including clerical and sales workers, routine workers, workers in agriculture, forestry and commercial fishing, manufacturing workers, distribution and service workers and other production workers.

Figure 6 plots the employment response for these three groups. The negative employment response is clearly concentrated among routine and manual workers. There is no substantial response for non-manual, non-routine lower-level workers and we observe an increase for upper-level workers. Table 4 confirms these observations and shows that there is a large negative effect for manual workers (-20.7 percent), and substantially smaller and noisier effects for lower-level employees (-7.4 percent), and upper-level employees (5.6 percent).

Investments. If the firm production function is such that capital and labor are substitutes, then the employment effects we estimate should result in an increase in investment to substitute for the decrease in labor. However, if capital and labor are complements, a decrease in employment should result in a decrease in investment. It is worth noting that this logic abstracts from any liquidity effects (firms not having enough cash to fund their operations): if the liquidity effects are larger than the substitution effects, an increase in the price of labor will lead to an increase in both capital and labor even if capital and labor are substitutes. We return to this point below, in Section 5.

Figure 7 shows the effect of the payroll tax discontinuity on investments. Panel a of Figure 7 shows that total investments decrease as a result of the higher payroll tax rates. This decrease in total investments is driven by a decrease in fixed asset investments (investment

[11] See also Akerman et al. [2015], Acemoglu and Restrepo [2018], Autor et al. [2003], Hershbein and Kahn [2018], Autor and Dorn [2013] and Goos et al. [2014].

in machines and equipment), as shown in panel b of Figure 7. We do not observe any change in buildings, as shown in panel c of Figure 7. And we observe an increase in R&D investment, as shown in panel d of Figure 7.

Table 5 provides the corresponding estimates. We estimate a decrease of 13.9 percent for total investment, which is mostly driven by an 18.0 percent decrease in fixed assets. Note that we also estimate a 24.4 percent increase in R&D investment, which only affects a few firms, since the majority do not invest in R&D. The corresponding placebo estimates show small and insignificant effects for total investment and all other investment subcategories.

Note that the mechanical positive correlation between investment and capital depreciation cannot explain these results for two main reasons. First, this correlation is positive, since the more a given firm invests, the more depreciation it claims, which would go against our findings that firms above the depreciation cutoff invest less. Second, this mechanical correlation should affect investment linearly, and should not create a discontinuity at the cutoff and therefore should not affect our estimation strategy.

Sales and productivity. Given that firms cut back on both capital and labor as a response to the increase in payroll taxes, one could reasonably expect a decrease in sales. The upper panel of Figure 8 plots the response of sales. The discontinuity at the threshold is negative. We estimate a response of -6.5 percent (relative to a placebo of -3.3 percent), implying that the volume of sales, while it responds negatively as one would expect, exhibits a limited response. Importantly, the sales response is also statistically indistinguishable from the placebo estimate.

We also estimate a large decrease in the use of intermediate inputs of -29.2 percent, which is consistent with the estimated decrease in sales. Note that, in principle, it could be that the employment response we estimate is due to outsourcing. However, outsourcing costs are included in inputs costs, and thus if firms were merely outsourcing employees, we should observe an increase in inputs costs instead of the estimated decrease.

The observed limited decrease in sales could be consistent with an increase in productivity, which would also be consistent with the fact that we estimate a decrease in less productive workers, i.e. manual and routine workers, as well as low-skilled workers. In Figure 8, we plot the response of labor productivity, which we define as the ratio of firm-level value added divided by labor costs. We find that labor productivity indeed increases at the threshold. In Table 6 we estimate that labor productivity increases by 12.8 percent, relative to a placebo estimate of -6.2 percent. Similarly, we observe that capital productivity – value added divided by annual investment – increases at the threshold by 11.3 percent and the placebo estimate is statistically insignificant at -5.9 percent.

These results suggest that firms could be mitigating the effects of the higher payroll tax rates by scaling down on the less productive factors of production.

Robustness checks. We perform the following five robustness checks. First, and as mentioned above, we systematically implement our estimates on the post-2010 period, when the discontinuity did not exist, as a placebo test and consistently estimate very small and statistically insignificant responses, which mitigates concerns over our identification strategy.

Second, we vary the bin size to ensure that the discontinuity observed in our graphical evidence is not driven by this choice. Appendix Figures 11 and 12 plot the responses of all our outcomes of interest using two smaller bin sizes compared to our standard choice. Changing the size of the bins does not affect our graphical evidence.

Third, we vary the size of the bandwidth in our estimations: Appendix Figure 13 shows the estimated employment and investment responses by size of bandwidth. While small bandwidths yield noisy estimates, the estimates stabilize relatively quickly and are virtually constant when considering any bandwidths, even those far from the optimal bandwidth derived from Calonico et al. [2014].

Fourth, we vary the size of the donut hole and re-estimate equation (1). Appendix Table 10 reports the estimated investment and employment responses by the size of the missing mass region (donut-hole region to the right of the threshold). Appendix Table 11 reports the estimated investment and employment responses by the size of the bunching region (donut-hole region to the left of the threshold). We estimate that the magnitude of the employment and investment responses is robust to varying the size of the donut hole, both to the right and to the left of the threshold.

Fifth, we re-estimate equation (1) using second- and third-degree polynomials. Appendix Table 12 reports the estimates for employment and investment. The estimates are also robust to the choice of polynomial degree, and, if anything, the magnitude of the estimates is larger with more flexible polynomials compared to our baseline estimates.

5 Implications for Firm-Level Production

Wages are downwards rigid. This is a finding that has been widely discussed in the labor economics literature.[12] Notably, Card [1990], shows that nominal wage rigidity leads to employment effects at the firm level. Yet, in spite of this large literature documenting the existence of wage rigidity, there is limited evidence of wages not responding to payroll tax increases, and the consensus in public finance is still that the incidence of payroll taxes is fully borne by workers.

In principle, in our setting, wages could be rigid because of the prevalence of collective sector-level wage agreements in Finland. These agreements set the industry-level minimum wage, but otherwise allow wages to vary flexibly, and therefore should only bind for low-earners. Therefore, they could explain why payroll taxes are fully borne by firms, but only for employees earning the collectively bargained minimum wage. In Appendix Table 9, we test whether the earnings of top earners respond to payroll taxes by showing the earnings responses by earnings deciles. We find that they do not, suggesting that collective wage agreements are not the explanation for payroll taxes being borne by firms, at least for non-

[12] See, for example, Akerlof et al. [1996], Kahn [1997], Card and Hyslop [1997], Dickens et al. [2007], Barattieri et al. [2014].

minimum wage-earning workers.

What explains downwards wage rigidity? While this is beyond the scope of this paper, several explanations have been put forward, including, more recently, explanations based on fairness and norms. Kaur [2019], for example, shows, using a survey of farm workers in India, that fairness considerations a la Kahneman et al. [1986] are likely to explain wage rigidity.

Interestingly, we find that belonging to a labor union has limited employment effects as shown in Figure 9 and Table 7. In other words, the employment responses seem to be very similar whether or not employees belong to a labor union or not. This could be due to two reasons: (1) labor unions tend to represent everyone in a particular industry, irrespective of whether an employee actually contributes, or (2) the employment effects are not driven by employees being fired but instead by fewer employees being hired. We believe both could be at play in our setting, and as the collective bargaining in wage-setting is widely applied across sectors in Finland, the first reason is likely to be very relevant.

At the micro level, we estimate that labor and capital are complements. When assuming a constant elasticity of substitution (CES) production function, our results imply a micro capital-labor elasticity of substitution that is equal to zero. We derive these predictions in Appendix Section C, but the intuition for this result is straightforward. If the elasticity of substitution between capital and labor is positive, then when labor decreases (after payroll taxes increase), capital should increase, as firms substitute away from labor towards capital. Instead, we estimate that, as labor decreases, so does capital, which implies that the two are complements in the CES framework and that the micro-level capital-labor elasticity of substitution is zero, i.e. that the production function is Leontief.

There are very few estimates of firm-level capital-labor elasticity of substitution. Two notable exceptions are Raval [2014] and Oberfield and Raval [2014], who estimate the capital-labor elasticity of substitution using micro data by relying on cross-sectional variation in local wages. Oberfield and Raval [2014] also offer a framework to aggregate micro elasticities into macro elasticities. Both papers estimate a capital-labor elasticity of substitution below one, but the estimates are well above zero.

Our paper provides one of the first of such estimates using a quasi-experimental setup. Our estimate is far from the macro estimate, and while Houthakker [1955] shows that even micro-level Leontief production functions can be aggregated to CES with a capital-labor elasticity of substitution greater than 1, we show, in Appendix Section C, using the aggregation framework from Oberfield and Raval [2014], that the macro-level capital-labor elasticity of substitution implied by our micro estimate is far smaller than 1. In principle, this could cast doubt on the argument put forth in Piketty [2014] that a fall in labor shares is likely driven by a capital-labor elasticity of substitution greater than 1.

However, if we step away from the CES framework, the positive correlation between capital and labor could also be consistent with liquidity effects being larger than substitution effects, which we explore below.

Liquidity Effects. Are liquidity constraints binding? In spite of this question being seemingly simple, there is no clear empirical answer to it. Modigliani and Miller [1958] predict that, with no differential costs of internal and external financing, firms should not face substantial liquidity constraints. On the other hand, if external financing is more costly than internal financing – possibly because of asymmetric information or incomplete contracting – cash injections should have a positive effect on capital expenditures. Rauh [2006], for example, uses a regression kink design at the pension funding threshold below which firms have to spend extra cash to ensure that their pensions are funded. He finds that the additional cash generated by the pension funding threshold affects capital expenditures but acknowledges the possibility that external financing costs might be discontinuously different above and below the pension funding cutoff, thus biasing the magnitude of the response upwards. Another example is Blanchard et al. [1994], who analyze the response of a sample of eleven firms to winning monetary payments from lawsuits and find no effect on capital expenditures, consistent with the prediction of Modigliani and Miller [1958]. Similarly, Saez et al. [2019] acknowledge that the firm-level effect of payroll taxes that they estimate is the combined effect on business activity of both cash windfalls and factor price changes. Because of their empirical design, which compares labor-intensive versus capital-intensive firms, they cannot disentangle these two effects.

Using our empirical design, we can investigate whether payroll taxes impose substantial liquidity constraints on firms by implementing a simple test: if labor and capital are complements, then we should observe a constant labor-to-capital ratio above and below the payroll tax discontinuity. Figure 10 shows a decrease in the labor-to-capital ratio as payroll taxes increase, implying that there could be substitution away from labor to capital that is masked overall by liquidity effects.

6 Conclusion

In this paper, we use quasi-experimental variation in payroll taxes to investigate how firms use their input factors. We uncover several new facts about firm behavior: as the cost of labor increases, (1) firms substitute away from low-skilled, routine and manual workers towards more productive workers, (2) firms decrease investments, while (3) productivity increases.

Our results have important implications for our understanding of firm-level production and input factor choices. First, our findings are inconsistent with large micro-level substitution between capital and labor, or at the very least larger than liquidity effects. Second, our results highlight the importance of accounting for heterogeneity in skill level and job tasks when estimating the incidence of payroll taxes. Third, from a policy perspective, our estimates imply that payroll taxes impose a negative fiscal externality on several other fiscal bases as they reduce capital but also profits. This effect should be taken into account when governments score payroll tax changes.

References

Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier, 2011.

Daron Acemoglu and Pascual Restrepo. The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542, 2018.

George A Akerlof, William T Dickens, George L Perry, Robert J Gordon, and N Gregory Mankiw. The macroeconomics of low inflation. *Brookings Papers on Economic Activity*, 1996(1):1–76, 1996.

Anders Akerman, Ingvil Gaarder, and Magne Mogstad. The skill complementarity of broadband internet. *The Quarterly Journal of Economics*, 130(4):1781–1824, 2015.

Pol Antras, Teresa C Fort, and Felix Tintelnot. The margins of global sourcing: theory and evidence from US firms. *American Economic Review*, 107(9):2514–64, 2017.

David Autor and David Dorn. The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–97, 2013.

David Autor, Lawrence Katz, and Melissa Kearney. The polarization of the US labor market. *American economic review*, 96(2):189–194, 2006.

David H Autor, Frank Levy, and Richard J Murnane. The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333, 2003.

Patrick Bajari, Han Hong, Minjung Park, and Robert Town. Regression discontinuity designs with an endogenous forcing variable and an application to contracting in health care. Working paper, National Bureau of Economic Research, 2011.

Alessandro Barattieri, Susanto Basu, and Peter Gottschalk. Some evidence on the importance of sticky wages. *American Economic Journal: Macroeconomics*, 6(1):70–101, 2014.

Alan I Barreca, Jason M Lindo, and Glen R Waddell. Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry*, 54(1):268–293, 2016.

Olivier Jean Blanchard, Florencio Lopez-de Silanes, and Andrei Shleifer. What do firms do with cash windfalls? *Journal of Financial Economics*, 36(3):337–360, 1994.

Antoine Bozio, Thomas Breda, and Julien Grenet. Does tax-benefit linkage matter for the incidence of social security contributions? Working paper, 2019.

Sebastian Calonico, Matias D Cattaneo, and Rocio Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.

David Card. Unexpected inflation, real wages, and employment determination in union contracts. *The American Economic Review*, pages 669–688, 1990.

David Card and Laura Giuliano. Does gifted education work? For which students? Working paper, National Bureau of Economic Research, 2014.

David Card and Dean Hyslop. Does inflation "grease the wheels of the labor market"? In *Reducing inflation: Motivation and strategy*, pages 71–122. University of Chicago Press, 1997.

David Card and Thomas Lemieux. Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, 116(2): 705–746, 2001.

Pedro Carneiro and Sokbae Lee. Trends in quality-adjusted skill premia in the United States, 1960-2000. *American Economic Review*, 101(6):2309–49, 2011.

Raj Chetty, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 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, 2011.

William T Dickens, Lorenz Goette, Erica L Groshen, Steinar Holden, Julian Messina, Mark E Schweitzer, Jarkko Turunen, and Melanie E Ward. How wages change: micro evidence from the international wage flexibility project. *Journal of Economic Perspectives*, 21(2): 195–214, 2007.

Claudia Goldin and Lawrence F Katz. The origins of technology-skill complementarity. *The Quarterly Journal of Economics*, 113(3):693–732, 1998.

Claudia Goldin and Lawrence F Katz. The race between education and technology: The evolution of US educational wage differentials, 1890 to 2005. Working paper, National Bureau of Economic Research, 2007.

Maarten Goos, Alan Manning, and Anna Salomons. Job polarization in Europe. *American Economic Review*, 99(2):58–63, 2009.

Maarten Goos, Alan Manning, and Anna Salomons. Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, 104(8):2509–26, 2014.

Brad Hershbein and Lisa B Kahn. Do recessions accelerate routine-biased technological change? Evidence from vacancy postings. *American Economic Review*, 108(7):1737–72, 2018.

Hendrik S Houthakker. The Pareto distribution and the Cobb-Douglas production function in activity analysis. *The Review of Economic Studies*, 23(1):27–31, 1955.

Shulamit Kahn. Evidence of nominal wage stickiness from microdata. *The American Economic Review*, 87(5):993–1008, 1997.

Daniel Kahneman, Jack L Knetsch, and Richard Thaler. Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, pages 728–741, 1986.

Lawrence F Katz and Kevin M Murphy. Changes in relative wages, 1963–1987: supply and demand factors. *The Quarterly Journal of Economics*, 107(1):35–78, 1992.

Supreet Kaur. Nominal wage rigidity in village labor markets. *American Economic Review*, 109(10):3585–3616, 2019.

Henrik J. Kleven and Mazhar Waseem. Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. *The Quarterly Journal of Economics*, 128(2):669–723, 2013. doi: 10.1093/qje/qjt004. URL http://qje.oxfordjournals.org/content/128/2/669.abstract.

Per Krusell, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, 68(5):1029–1053, 2000.

H Ku, U Schoenberg, and R Schreiner. Do place-based tax incentives create jobs? *Journal of Public Economics*, 2020.

Andreas Lichter, Andreas Peichl, and Sebastian Siegloch. The own-wage elasticity of labor demand: A meta-regression analysis. *European Economic Review*, 80:94–119, 2015.

Franco Modigliani and Merton H Miller. The cost of capital, corporation finance and the theory of investment. *The American Economic Review*, 1:3, 1958.

Ezra Oberfield and Devesh Raval. Micro data and macro technology. 2014.

Thomas Piketty. *Capital in the 21st Century*. Harvard University Press Cambridge, MA, 2014.

Joshua D Rauh. Investment and financing constraints: Evidence from the funding of corporate pension plans. *The Journal of Finance*, 61(1):33–71, 2006.

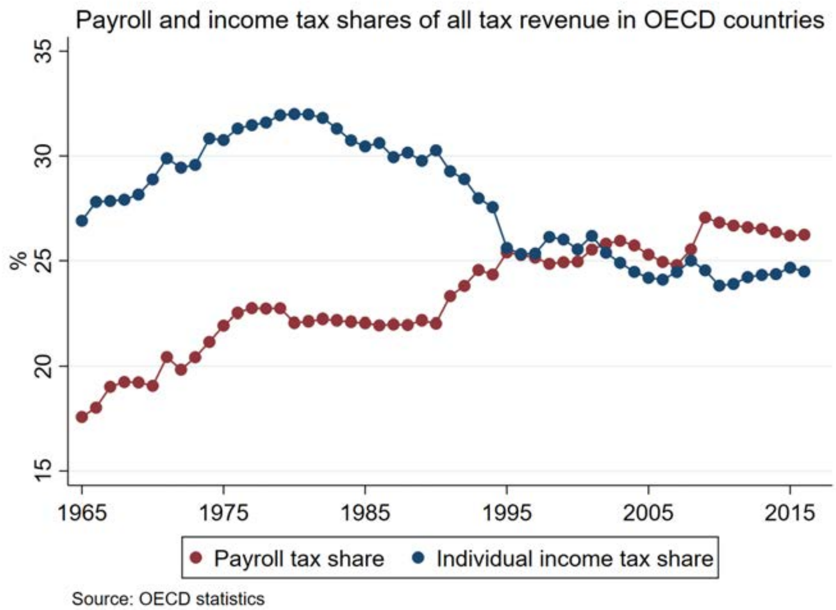
Devesh Raval. The micro elasticity of substitution and non-neutral technology. Working paper, 2014.

Emmanuel Saez, Manos Matsaganis, and Panos Tsakloglou. Earnings determination and taxes: evidence from a cohort-based payroll tax reform in Greece. *The Quarterly Journal of Economics*, 127(1):493–533, 2012.

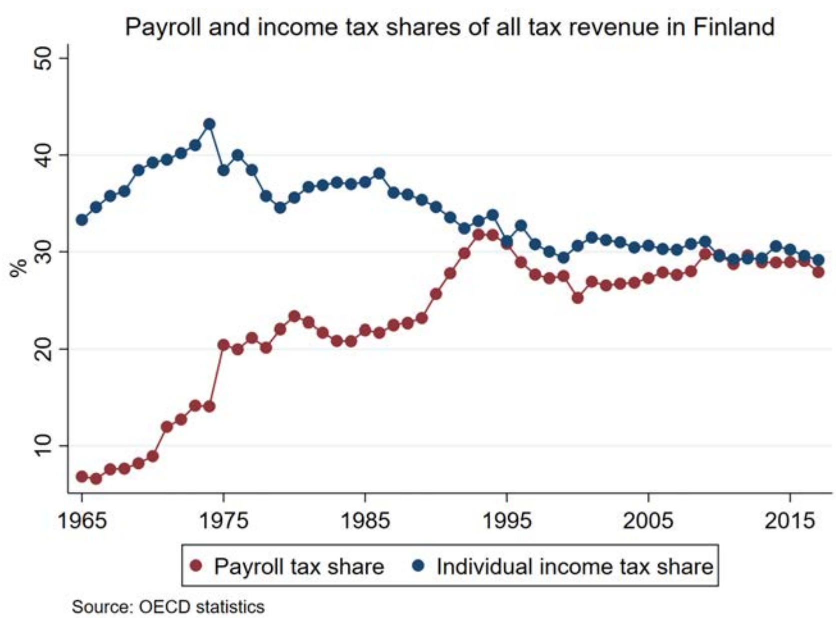
Emmanuel Saez, Benjamin Schoefer, and David Seim. Payroll taxes, firm behavior, and rent sharing: Evidence from a young workers' tax cut in Sweden. *American Economic Review*, 109(5):1717–63, 2019.

Figure 1: Payroll and Income Tax Shares of Total Tax Revenue

(a) OECD Countries



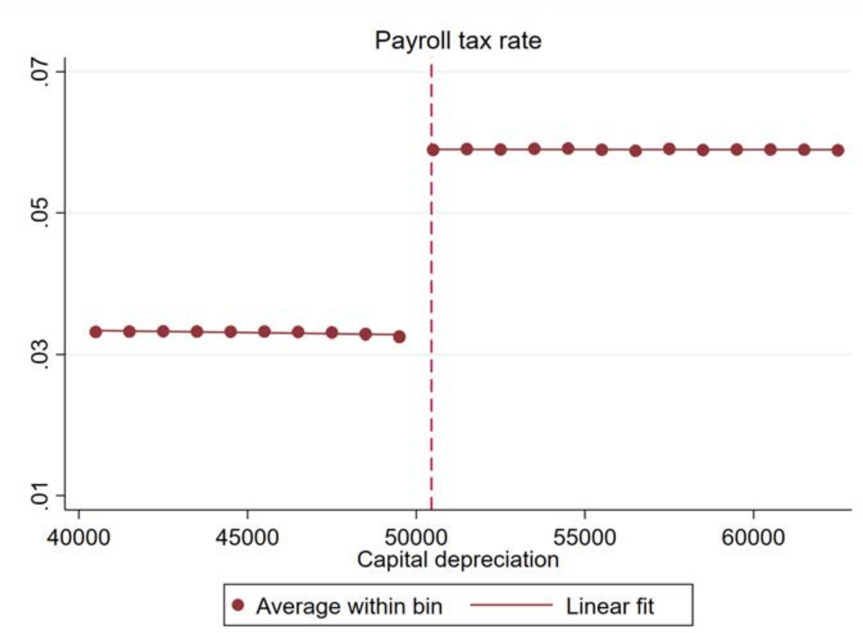
(b) Finland



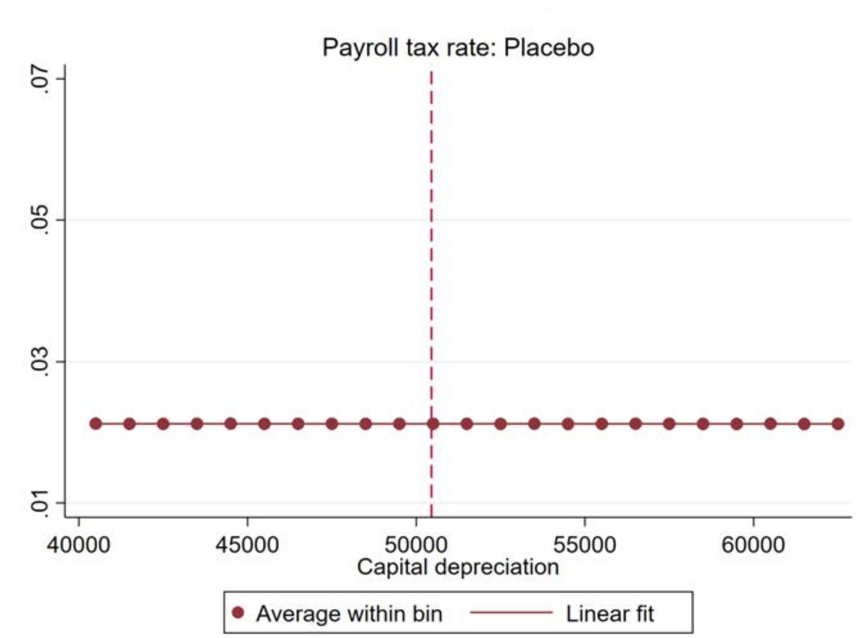
Notes: These Figures plot the share of total tax revenue raised by payroll and income taxes over time in the OECD countries (Figure 1a) and in Finland (Figure 1b).

Figure 2: Payroll Tax Rate

(a) Treatment Years (1996-2009)



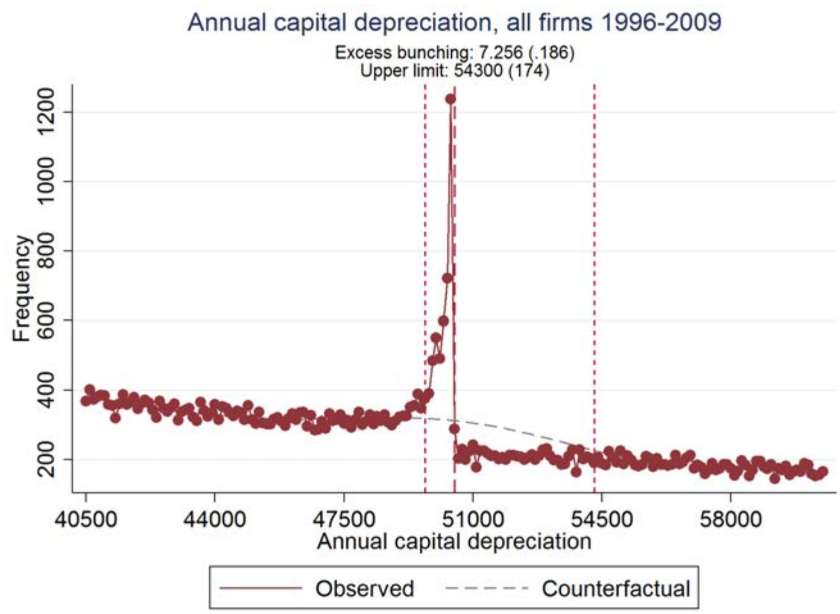
(b) Placebo Years (2010-2015)



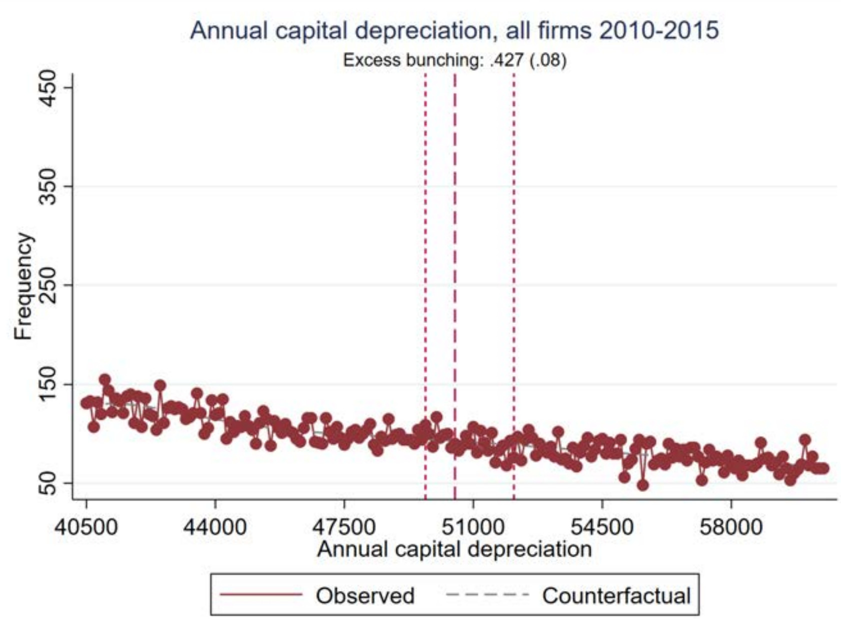
Notes: This Figure plots the average payroll tax rates above and below the capital depreciation threshold for health and pension contributions. The second panel shows a placebo test for years 2010 to 2015 for the same variable.

Figure 3: Distribution Around Cutoff

(a) Treatment Years (1996-2009)



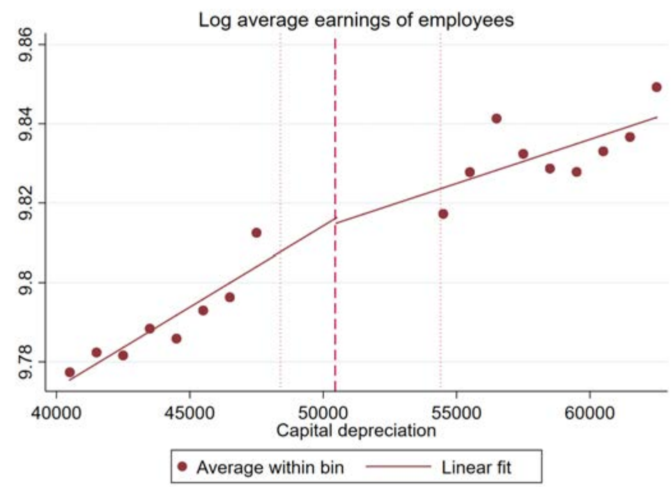
(b) Placebo Years (2010-2015)



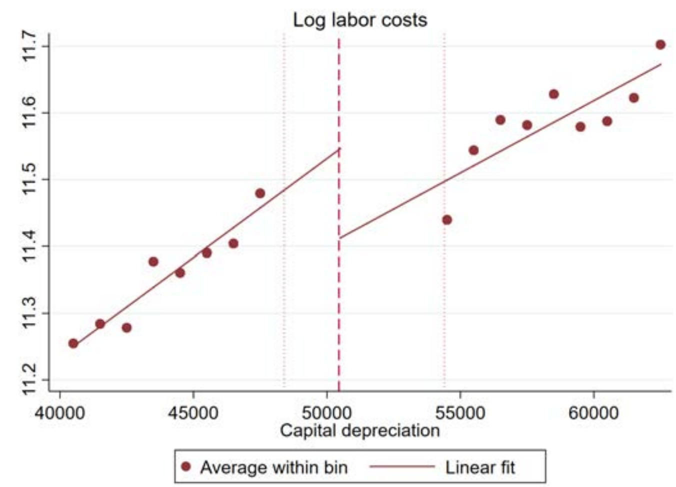
Notes: These Figures plot the distributions of capital depreciation in the treatment years (1996-2009) and placebo years (2010-2015) around the threshold. We follow the methods of Kleven and Waseem [2013] to estimate the excess mass at the threshold and determine the manipulated area, corresponding to the area of the excess and the missing masses. This approach is explained in more detail in Appendix Section B.

Figure 4: Earnings and Labor Costs

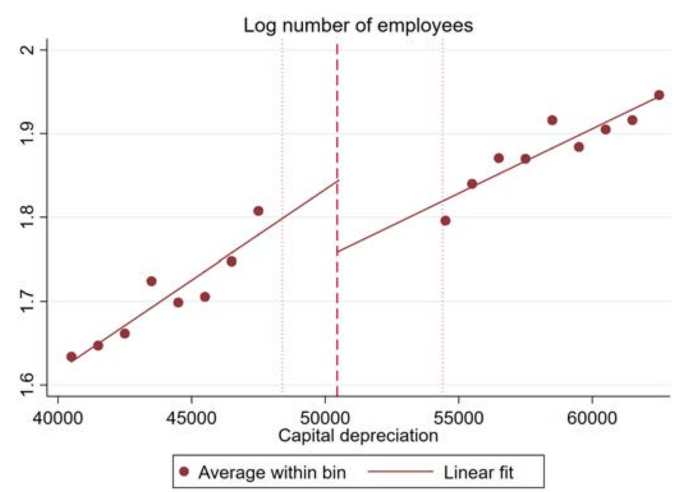
(a) Earnings



(b) Labor Costs

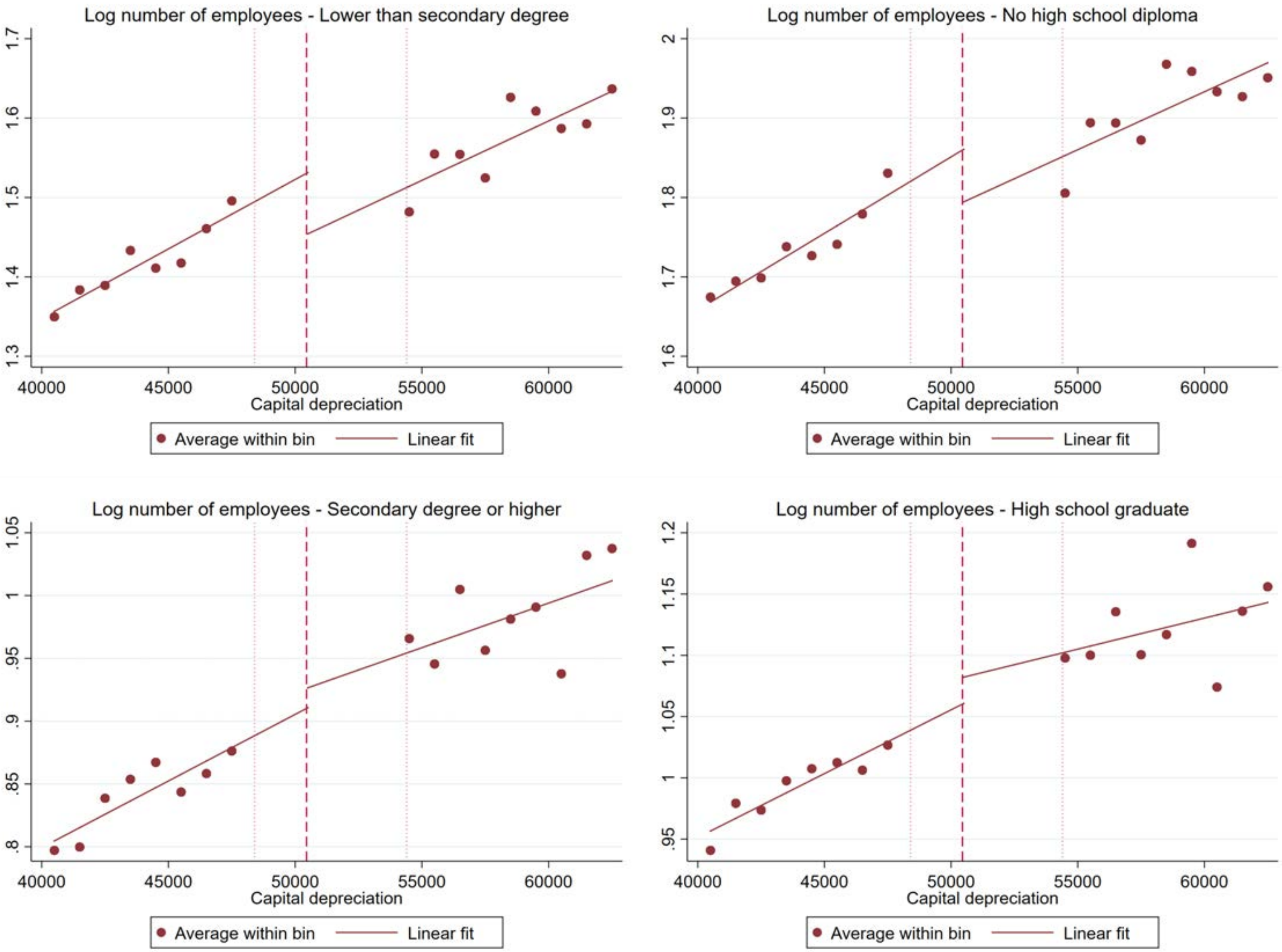


(c) Number of Employees



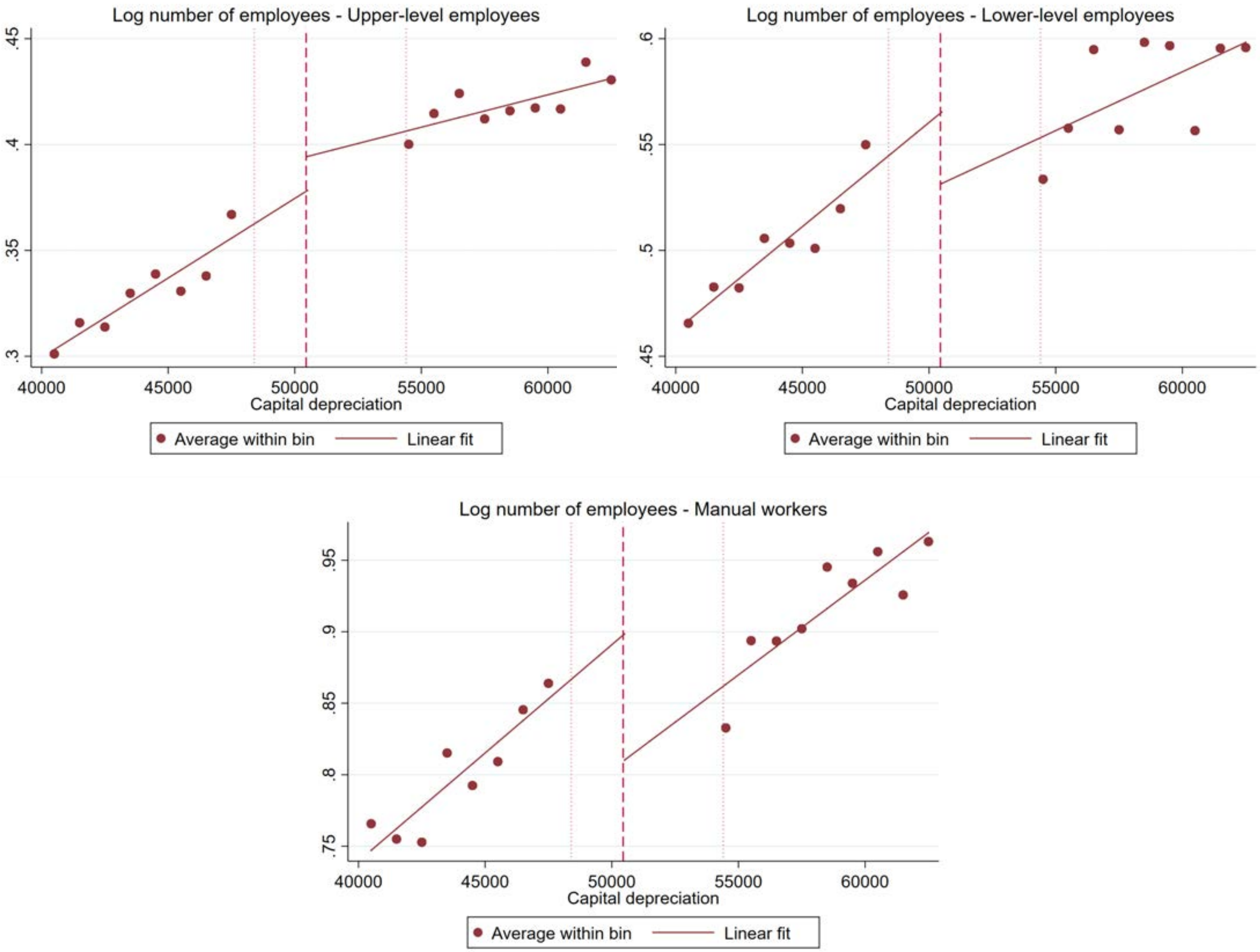
Notes: The first panel shows the response of earnings per employee (in logs) at the payroll tax discontinuity. The second panel shows the response of labor costs net of payroll taxes paid by firms to the payroll tax discontinuity. The years included are 1996 to 2009.

Figure 5: Employment Effects by Skill Level



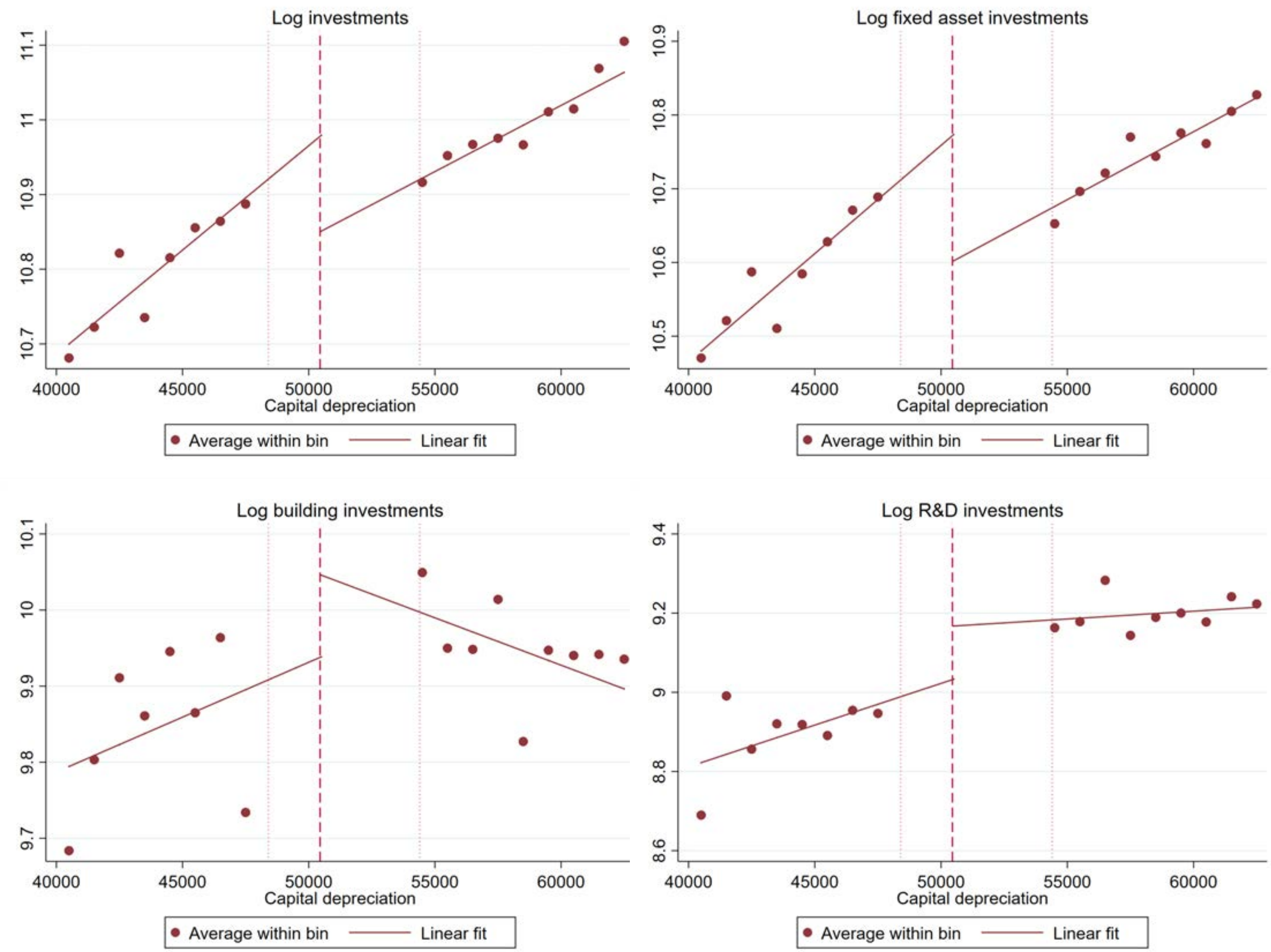
Notes: These Figures plot the (log) number of employees with lower than secondary education (first panel), with no high school diploma (second panel), with higher than secondary education (third panel) and with a high school degree in firms around the capital depreciation threshold.

Figure 6: Employment Effects by Job Task



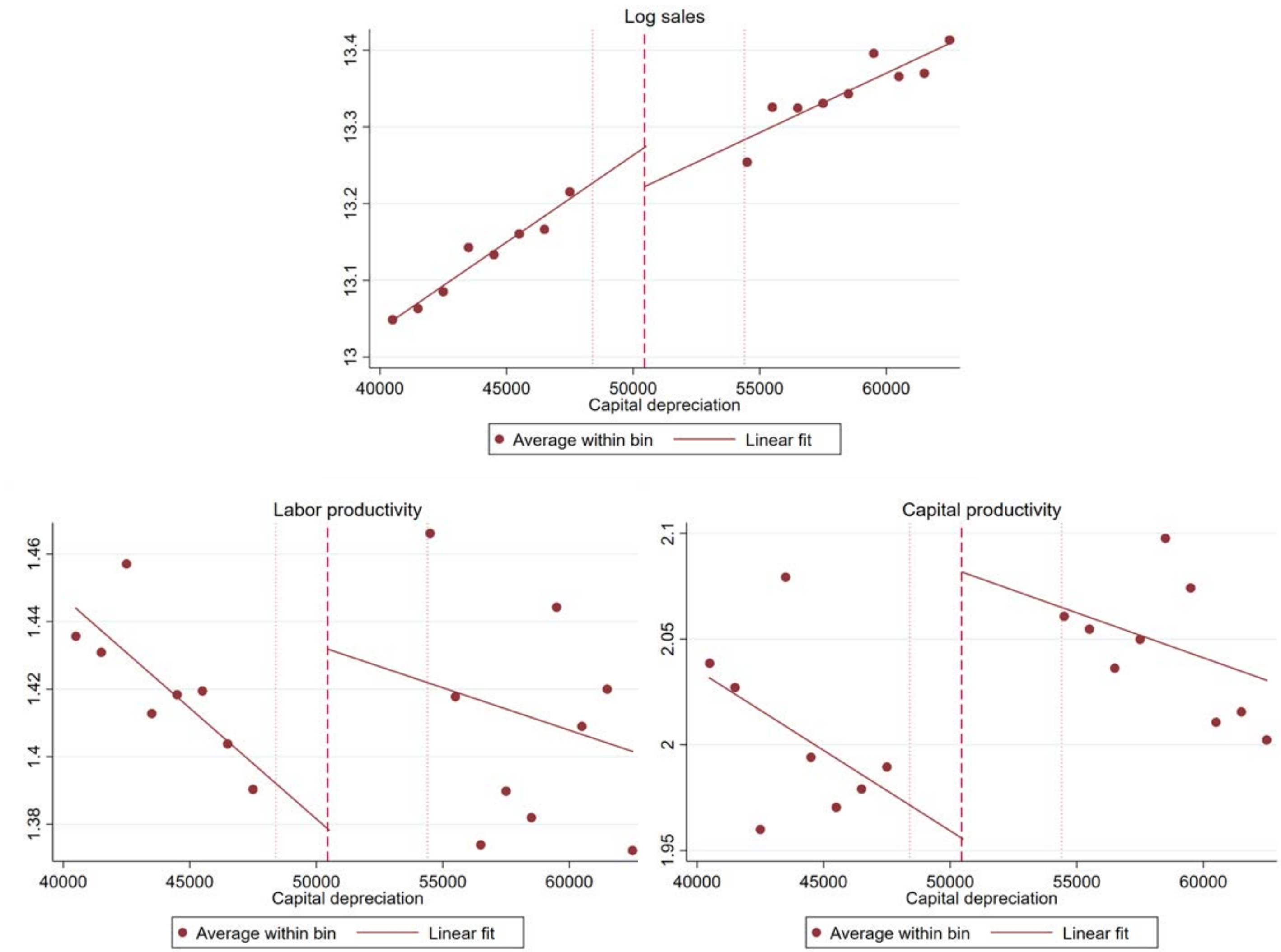
Notes: These Figures plot the (log) number of employees by type of task.

Figure 7: Effects on Investments



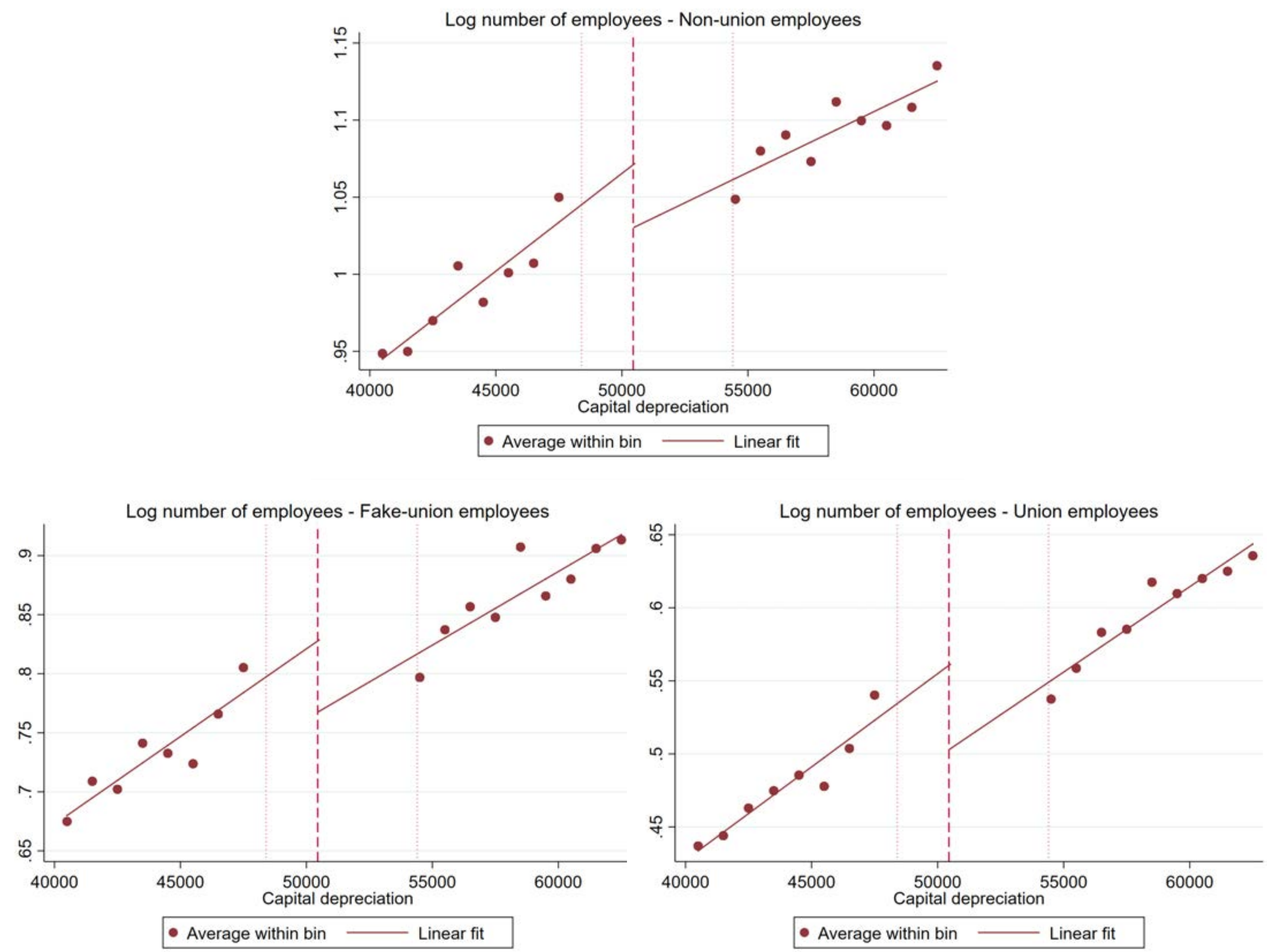
Notes: These Figures plot the (log) total annual investments of firms (first panel), and total investments divided by fixed assets (second panel), buildings (third panel) and R&D (fourth panel) around the capital depreciation threshold.

Figure 8: Production and Productivity



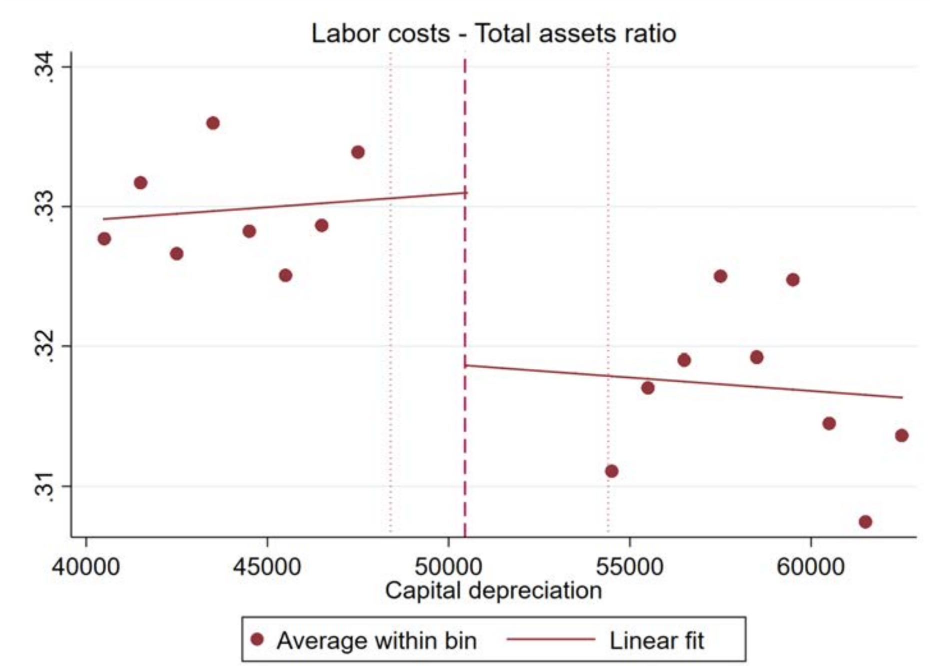
Notes: These Figures plot the (log) sales (first panel), labor (lower-left panel) and capital productivity (lower-right panel) of firms around the capital depreciation threshold.

Figure 9: Employment by Unionization Status



Notes: These Figures plot the (log) number of non-union employees, employees paying unemployment insurance payments but not belonging to a union (fake union) and employees belonging to a labor union around the capital depreciation threshold.

Figure 10: Liquidity Test



Notes: This Figure plots the share of labor costs divided by total assets around the capital depreciation threshold.

Table 2: Effects on Earnings and Number of Employees

Outcomes (logs)	Earnings	Labor Costs	No. Employees
		Treatment	
RD Estimate	-0.003	-0.175***	-0.089***
	(0.008)	(0.041)	(0.028)
Bandwidth	11,417	17,540	19,011
N above	12,832	22,749	24,115
N below	29,299	61,477	67,368
		Placebo	
RD Estimate	-0.032	0.076	0.034
	(0.023)	(0.063)	(0.037)
Bandwidth	8,773	14,237	18,214
N above	5,438	9,865	12,771
N below	10,491	21,308	29,923

Notes: This Table reports the results of estimating equation (1) on log earnings (first column), log labor costs (second column) and log number of employees (third column). The upper panel shows the results for the pre-2010 period (treatment) and the bottom panel shows the results for the post-2010 period (placebo). The Table shows the bias-corrected estimates with robust standard errors, the size of the optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within the bandwidth, respectively, following Calonico et al. [2014].

Table 3: Effects by Skills

Log No. Employees	High School	No Highs School	Secondary or Higher	Lower than Secondary
		Treatment		
RD Estimate	0.047	-0.221***	0.055	-0.169***
	(0.056)	(0.063)	(0.049)	(0.060)
Bandwidth	11,626	8,241	12,253	8,709
N above	8,567	7,811	8,204	8,792
N below	18,505	18,749	19,091	18,784
		Placebo		
RD Estimate	-0.081	0.008	-0.076	-0.036
	(0.149)	(0.052)	(0.112)	(0.061)
Bandwidth	6,851	11,623	7,571	10,530
N above	2,336	7,812	2,608	6,666
N below	4,577	15,402	4,777	12,734

Notes: This Table reports the results of estimating equation (1) on log number of employees with a high school degree (first column), with no high school degree (second column) with a college degree or higher (third column) and without a college degree (forth column). The upper panel shows the results for the pre-2010 period (treatment) and the bottom panel shows the results for the post-2010 period (placebo). The Table shows the bias-corrected estimates with robust standard errors, the size of the optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within the bandwidth, respectively, following Calonico et al. [2014].

Table 4: Effects by Job Tasks

Log No. Employees	Upper-level workers	Lower-level workers	Manual workers
		Treatment	
RD Estimate	0.056*	-0.074*	-0.207***
	(0.029)	(0.041)	(0.047)
Bandwidth	14,997	11,700	10,496
N above	20,100	14,908	12,901
N below	49,474	35,022	30,261
		Placebo	
RD Estimate	0.030	0.022	-0.005
	(0.042)	(0.051)	(0.050)
Bandwidth	10,505	11,801	12,612
N above	7,817	9,046	9,781
N below	15,489	18,190	19,951

Notes: This Table reports the results of estimating equation (1) on the number of upper-level employees (first column), log lower-level employees (second column) and log manual workers (third column). The upper panel shows the results for the pre-2010 period (treatment) and the bottom panel shows the results for the post-2010 period (placebo). The Table shows the bias-corrected estimates with robust standard errors, the size of the optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within the bandwidth, respectively, following Calonico et al. [2014].

Table 5: Effect on Investments

Outcomes (logs)	Investment	Fixed assets	Buildings	R&D
		Treatment		
RD Estimate	-0.139***	-0.180***	0.120	0.244**
	(0.043)	(0.047)	(0.119)	(0.101)
Bandwidth	20,376	18,755	13,284	15,799
N above	24,993	22,227	7,514	10,006
N below	71,862	61,295	17,493	22,601
		Placebo		
RD Estimate	0.052	0.016	0.128	0.158
	(0.066)	(0.067)	(0.150)	(0.165)
Bandwidth	16,538	16,535	14,402	12,906
N above	11,144	10,536	4,918	4,057
N below	25,139	23,743	9,969	7,762

Notes: This Table reports the results of estimating equation (1) on log total investments (first column), log investment in fixed assets (second column), log investment in buildings (third column) and log investment in research and development (fourth column). The upper panel shows the results for the pre-2010 period (treatment) and the bottom panel shows the results for the post-2010 period (placebo). The Table shows the bias-corrected estimates with robust standard errors, the size of the optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within the bandwidth, respectively, following Calonico et al. [2014].

Table 6: Effect on Firm Production Measures

Outcomes (logs)	Sales	Inputs	Markup	Labor Productivity	Capital Productivity
		Treatment			
RD Estimate	-0.065**	-0.292**	0.014	0.128***	0.113**
	(0.027)	(0.128)	(0.054)	(0.028)	(0.046)
Bandwidth	25,227	27,884	23,658	19,224	23,112
N above	32,834	36,345	22,840	24,615	27,615
N below	114,189	118,055	74,573	69,495	86,439
		Placebo			
RD Estimate	-0.033	0.142	-0.008	-0.062	-0.059
	(0.050)	(0.163)	(0.064)	(0.038)	(0.089)
Bandwidth	15,997	14,461	18,405	14,672	14,868
N above	12,225	11,121	13,219	10,110	9,813
N below	28,483	24,964	32,494	21,956	21,345

Notes: This Table reports the results of estimating equation (1) on log sales (first column), log intermediate inputs (second column), log markup (third column), labor productivity (fourth column) and capital productivity (fifth panel). The upper panel shows the results for the pre-2010 period (treatment) and the bottom panel shows the results for the post-2010 period (placebo). Table shows bias-corrected estimates with robust standard errors, the size of optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within bandwidth, respectively, following Calonico et al. [2014].

Table 7: Role of Unions

Outcomes	Share of union employees	No. not union employees	No. fake union employees	No. union employees
		Treatment		
RD Estimate	-0.014	-0.056*	-0.138***	-0.107***
	(0.015)	(0.030)	(0.044)	(0.036)
Bandwidth	9,372	14,527	9,456	9,860
N above	8,966	19,418	11,114	11,805
N below	20,343	47,297	26,246	27,809
		Placebo		
RD Estimate	0.001	0.052	0.070	-0.015
	(0.016)	(0.043)	(0.045)	(0.033)
Bandwidth	10,951	12,101	11,213	11,030
N above	6,560	9,327	8,479	8,304
N below	12,131	18,818	16,935	16,522

Note: This Table reports the results of estimating equation (1) on log share of union employes (first column), log number of employees not affiliated with a union (second column), log number of employees affiliated with a union that provides insurance benefits but no representation (fake union) (third column) and log number of employees affiliated with a union that provides insurance benefits and representation (fourth column). See Table 17 in Appendix for exact definitions for fake union and real union measures. The upper panel shows the results for the pre-2010 period (treatment) and the bottom panel shows the results for the post-2010 period (placebo). Table shows bias-corrected estimates with robust standard errors, the size of optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within bandwidth, respectively, following Calonico et al. [2014].

ONLINE APPENDIX NOT FOR PUBLICATION

A Depreciation Rules

The Finnish tax authorities' definition of capital is any fixed assets, including all long-term tangibles that firms use in their production process to generate income that cannot easily be converted into cash, such as land, buildings, machinery, stocks, equipment, vehicles, leasehold improvements, and other such items. Firms can choose their depreciation rules: (1) linear depreciation with the same euro value per year, or (2) double declining balance depreciation with the same percentage per year. In Finland, buildings, other constructions, machinery and equipment are all depreciated using the declining balance method. There are also different depreciation rules and percentages for different asset types. Depreciation for each building is calculated separately, with the maximum depreciation percentage varying from 4 percent to 20 percent, depending on the type of construction. For example, the annual depreciation rate for office buildings is 4 percent, 7 percent for factory buildings and 20 percent for immovable capital. The maximum rate of depreciation of machinery and equipment is 25 percent.

The life of assets can vary depending on the asset type and this directly affects the amount of depreciation. Assets with a useful life of less than three years may be written off using the free depreciation method, i.e. deduct up to 100 percent of the cost of assets in a single tax year where the value of each item is less than 850 euros and the total value of such assets is no more than 2,500 euros per tax year. Patents and other intangible rights, such as goodwill, are amortized on a straight-line basis for ten years, unless the taxpayer demonstrates that the asset has a shorter useful life.

B Bunching Methodology

We follow Chetty et al. [2011] and Kleven and Waseem [2013] in estimating the magnitude of bunching. First, we construct the counterfactual density by excluding the "distorted distribution" close to the observed distribution, and then fit a flexible polynomial function using the undistorted distribution.

We begin by constructing a bin sample. We divide the data into 100 euro bins and count the number of firms in each bin. Then we estimate a counterfactual density by running the following regression while excluding the region around the threshold $[D_L, D_H]$:

$$c_j = \sum_{i=0}^{p} \beta_i (D_j)^i + \sum_{i=D_L}^{D_H} \eta_i \cdot \mathbf{1}(D_j = i) + \varepsilon_j \tag{2}$$

where c_j is the count of firms in bin j, D_j denotes the depreciation in bin j and p is the order of the polynomial. Therefore, the estimated values for the counterfactual density are $\hat{c}_j = \sum_{i=0}^{p} \beta_i (D_j)^i$. We can calculate the excess bunching by comparing the actual number of firms just below the threshold (within (D_L, D^*)) to the estimated counterfactual density within the same region:

$$\hat{b}(D^*) = \frac{\sum_{i=D_L}^{D^*} (c_j - \hat{c}_j)}{\sum_{i=D_L}^{D^*} \hat{c}_j / N_j}$$

where N_j represents the number of bins within $[D_L, D^*]$.

As is common in the bunching literature, we define the lower limit of the excluded region (D_L) simply based on visual observations, representing the point where bunching begins.

We follow the approach of Kleven and Waseem [2013] to define the upper limit and thus the marginal buncher firm D_H. This point is determined such that the estimated excess mass equals the estimated missing mass above the threshold D^*. In practice, we do this using an iterative process which starts with a small D_H and converges when the excess mass is equal to the missing mass, i.e. $\hat{b}_E(y^*) \approx \hat{b}_M(y^*)$.

Finally, we calculate standard errors by using a residual-based bootstrap procedure. We first generate a large number of depreciation distributions by randomly resampling the residuals from equation (2) with a replacement. Then, based on the resampled distributions, we estimate a large number of counterfactual densities. In the bootstrap procedure, we also take into account the iterative process to determine the marginal buncher. Based on these bootstrapped counterfactual densities, we evaluate variation in the estimates of interest. The standard errors for each estimate are defined as the standard deviation in the distribution of the estimate.

C Capital-Labor Elasticity of Substitution: Conceptual Framework

C.1 Micro Capital-Labor Elasticity of Substitution

Production Function. We assume that firms exhibit constant elasticity of substitution (CES) production functions as follows:

$$F(k,l) = (\alpha k^{\frac{\sigma-1}{\sigma}} + (1-\alpha) l^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}},$$

where k is capital, l is labor, and α and σ are parameters. σ is assumed to be strictly positive and has no upper bound. When $\sigma \to 0$, it can be shown that the production function is Leontief with the following form:

$$F(k,l) = \min(k,l).$$

Denote by $\epsilon_{k,l}$ the elasticity of substitution between capital and labor and by RTS the rate of technical substitution between capital and labor. It can be shown that the capital-labor substitution elasticity only depends on σ:

$$\epsilon_{k,l} = \frac{d(k/l)}{d(RTS)} \frac{RTS}{k/l} = \frac{d(k/l)}{d(-F_l/F_k)} \frac{-F_l/F_k}{k/l} = \sigma.$$

Next, since we are interested in how capital and labor respond to changes in payroll taxes, we derive the demands for labor and capital by minimizing the cost function subject to a production level constraint. We assume $\sigma > 0$ throughout and return to Leontief production functions below. Formally, we solve the following minimization problem for $\sigma > 0$, where w is wages and r is the cost of capital:

$$\min_{k,l} C(w,r) = wl + rk$$

subject to

$$F(k,l) = q_0$$

This yields the following condition:

$$k = \left(\frac{w}{r} \frac{\alpha}{1-\alpha}\right)^{\sigma} l$$

Using this relationship and the resource constraint $F(k,l) = q_0$, we get:

$$l = q_0 \left(\alpha \left(\frac{w}{r} \frac{\alpha}{1-\alpha}\right)^{\sigma-1} + (1-\alpha)\right)^{\frac{\sigma}{1-\sigma}},$$

$$k = q_0 \left((1-\alpha) \left(\frac{w}{r} \frac{\alpha}{1-\alpha}\right)^{1-\sigma} + \alpha\right)^{\frac{\sigma}{1-\sigma}}.$$

We take the derivative of these two equations with respect to w to get the elasticity of capital and labor with respect to wages:

$$\epsilon_{k,w} = \frac{\partial k}{\partial w}\frac{w}{k} = \frac{(1-\alpha)\sigma}{(1-\alpha)+\alpha(\frac{w}{r}\frac{\alpha}{1-\alpha})^{\sigma-1}},$$

$$\epsilon_{l,w} = \frac{\partial l}{\partial w}\frac{w}{l} = -\frac{\alpha\sigma}{\alpha+(1-\alpha)(\frac{w}{r}\frac{\alpha}{1-\alpha})^{1-\sigma}}.$$

These two expressions imply that firms with CES production functions with $\sigma > 0$ will increase capital when wages decrease and decrease labor when wages increase. Empirically, firms with CES production functions would respond to labor cost changes by decreasing their number of employees and increasing their capital investment to replace workers.

Leontief Production Function. Leontief production functions are a special case of CES production functions: it can be shown that when $\sigma \to 0$, i.e. the capital-labor supply elasticity tends to zero, which means that capital cannot be substituted with labor and vice-versa, $F(k,l) = \min(\alpha k, \beta l)$. In this case, labor and capital are used in equal shares. For this reason, when the cost of labor increases, both the demand for labor and for capital decrease. This implies that when the capital-labor elasticity of substitution is zero, both $\epsilon_{k,w}$ and $\epsilon_{l,w}$ will be negative. Empirically, when labor costs increase, firms with Leontief production functions reduce both their number of employees and their investment in capital since both inputs are used in fixed proportions.

A Simple Empirical Test of Leontief versus CES Production Functions. The derivations above imply a simple test of whether $\epsilon_{k,l}$ is strictly positive or zero: estimating the response of capital flows, i.e. investments, to labor cost changes. If investments *increase* when labor costs increase, then $\epsilon_{k,l} > 0$. If instead, investments *decrease* when labor costs increase then $\epsilon_{k,l} = 0$. In the rest of the paper, we set up our empirical framework to estimate how investments respond to changes in labor costs.

D Macro Elasticities

The capital-labor elasticity of substitution we have estimated is a micro elasticity and does not account for possible substitution across different firms andor industries. However, we can use our micro elasticity to derive an estimate of the macro elasticity by relying on the framework of Oberfield and Raval [2014]. The authors show that the aggregate elasticity of substitution is a weighted average of the micro elasticity of substitution and the elasticity of demand.

Formally, given the following production function: $F(k,l) = (\alpha k^{\frac{\sigma-1}{\sigma}} + (1-\alpha)l^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}}$, we denote by $\alpha_i = \frac{rk_i}{rk_i+wl_i}$ and $\alpha = \frac{rk}{rk+wl}$ the capital share in the total costs of production for firm i and the aggregate capital share, respectively. Further, we define θ_i to be plant i's cost of labor and capital as a share of the aggregate costs of labor and capital. Oberfield and Raval [2014] show that the macro capital-labor elasticity of substitution σ^{agg} is a weighted average of the micro elasticity of substitution and the elasticity of demand ε:

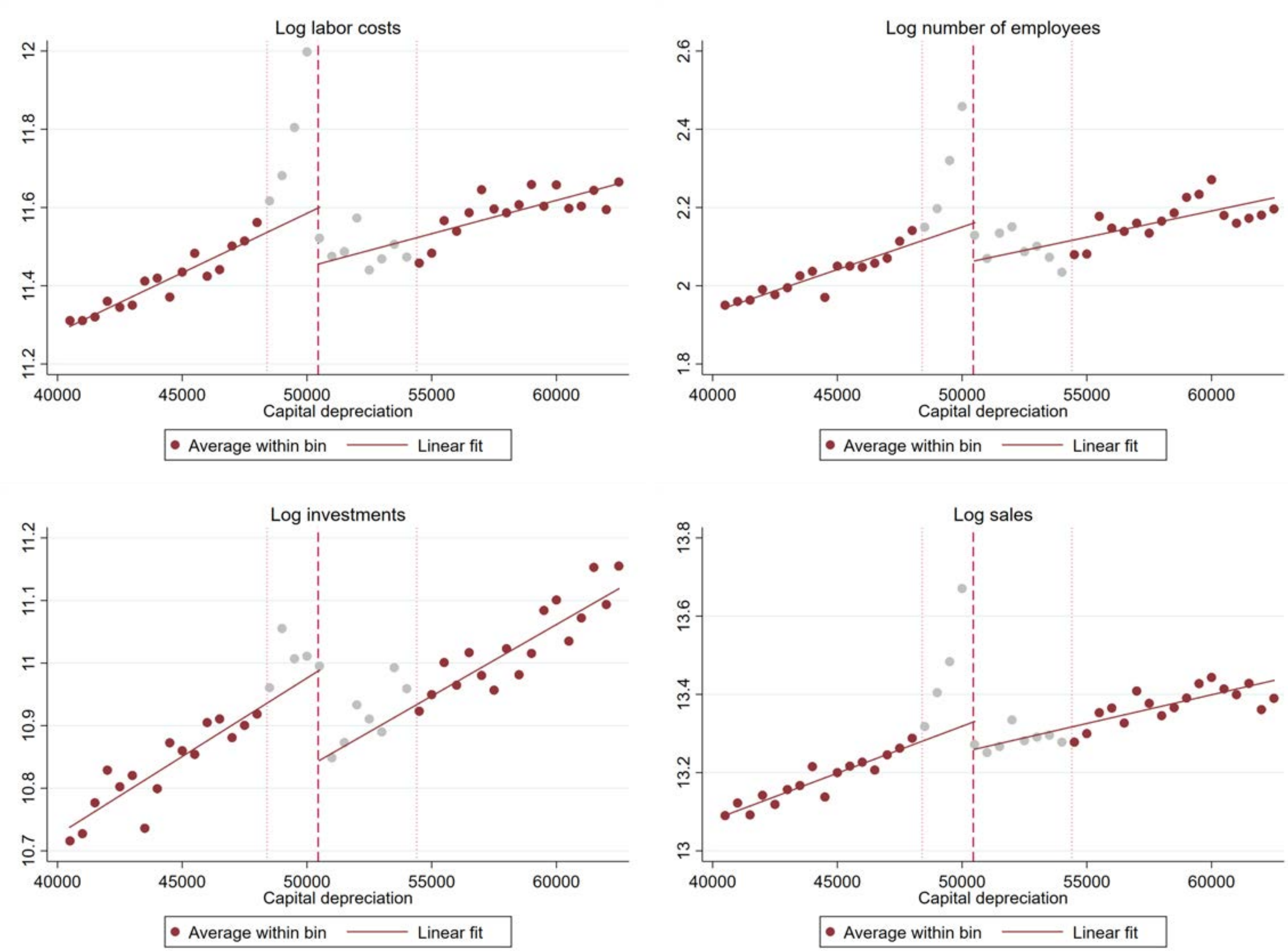
$$\forall \sigma \geq 0,\ \sigma^{agg} = (1-\chi)\sigma + \chi\varepsilon \tag{3}$$

where $\chi = \sum_{i \in I} \frac{(\alpha_i-\alpha)^2}{\alpha(1-\alpha)}\theta_i$ represents the degree of heterogeneity in the relative use of labor and capital in a given market and I is the total number of firms. $(1-\chi)\sigma$ measures the substitution of labor with capital within a given plant as a response to changes in relative factor prices and $\chi\varepsilon$ measures the reallocation effect of labor and capital across firms when relative factor prices change: for example, when the cost of capital increases, firms that rely more heavily on labor gain a cost advantage that they can pass through to prices. The elasticity of demand ε determines the extent to which consumers respond to lower prices by shifting consumption to the labor-intensive commodity.

α_i, α and θ_i are directly observable in the corporate tax data, which report both labor and capital costs. To estimate ε, we use the average markup μ and assume that $\varepsilon = 1/\mu$. We follow Antras et al. [2017] and define markups as $\frac{sales-costs}{costs}$.

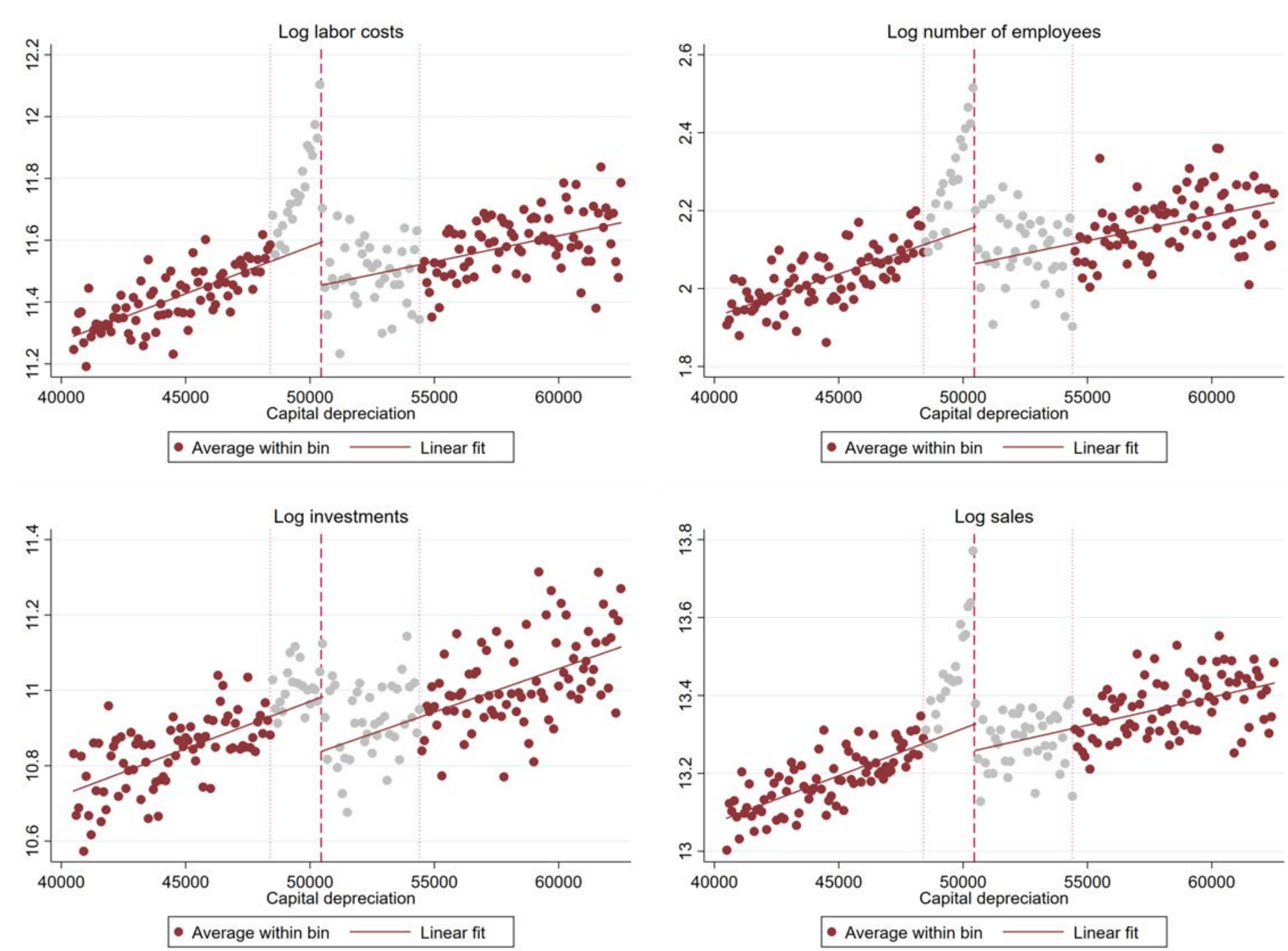
We estimate that $\chi = 0.13$ and $\varepsilon = 1.29$. These estimates imply a macro capital-labor elasticity of substitution $\sigma^{agg} = 0.17$.

Figure 11: Smaller Bin Width: 500 euros



Notes: These Figures show the main firm-level outcomes with a smaller, 500 euro, bin width at the payroll tax discontinuity from 1996 to 2009. In these Figures, we also plot the mean outcomes within the donut hole region.

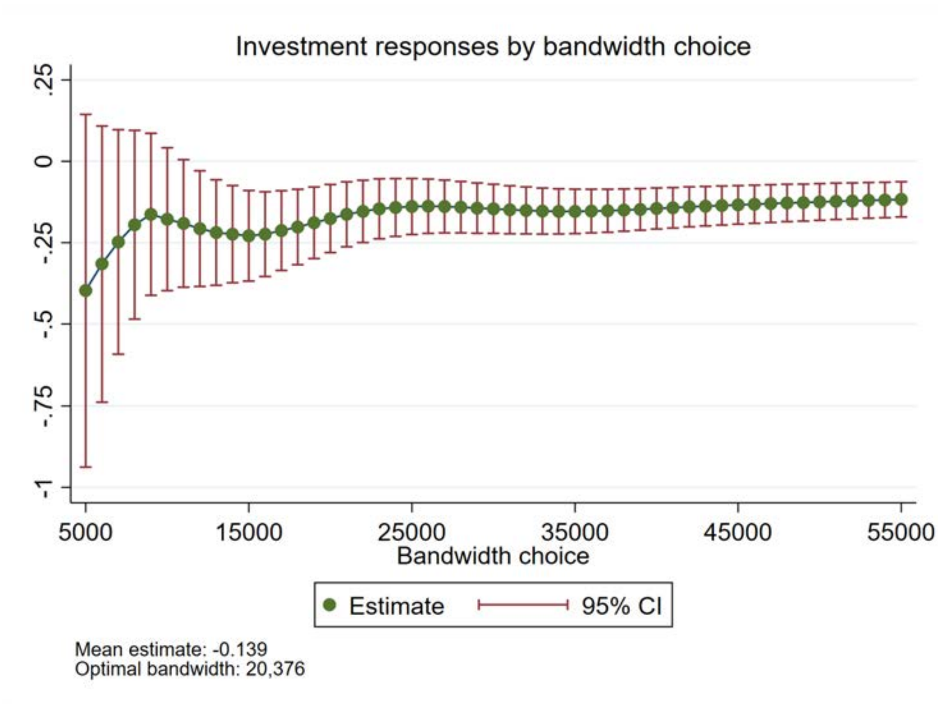
Figure 12: Smaller Bin Width: 100 euros



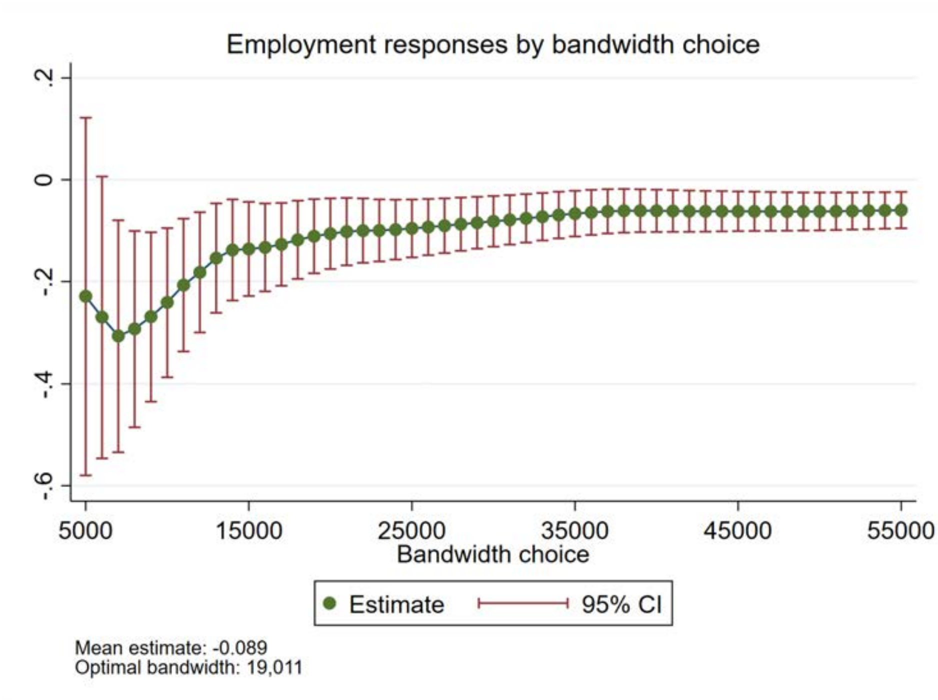
Notes: These Figures show the main firm-level outcomes with a smaller, 100 euro, bin width at the payroll tax discontinuity from 1996 to 2009. In these Figures, we also plot the mean outcomes within the donut hole region.

Figure 13: Estimates by Different Bandwidth

(a) Investment Response by Bandwidth

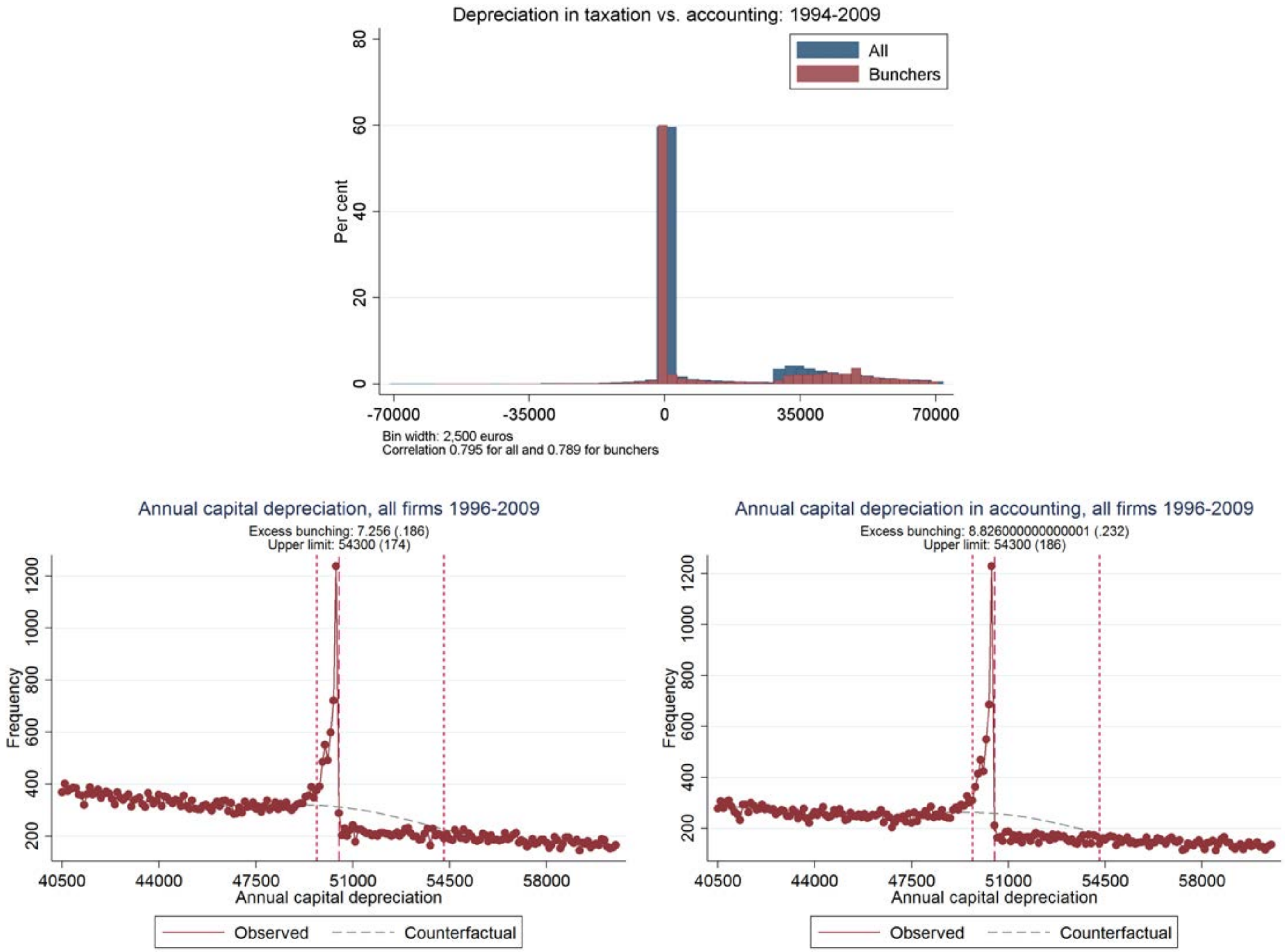


(b) Employment Response by Bandwidth



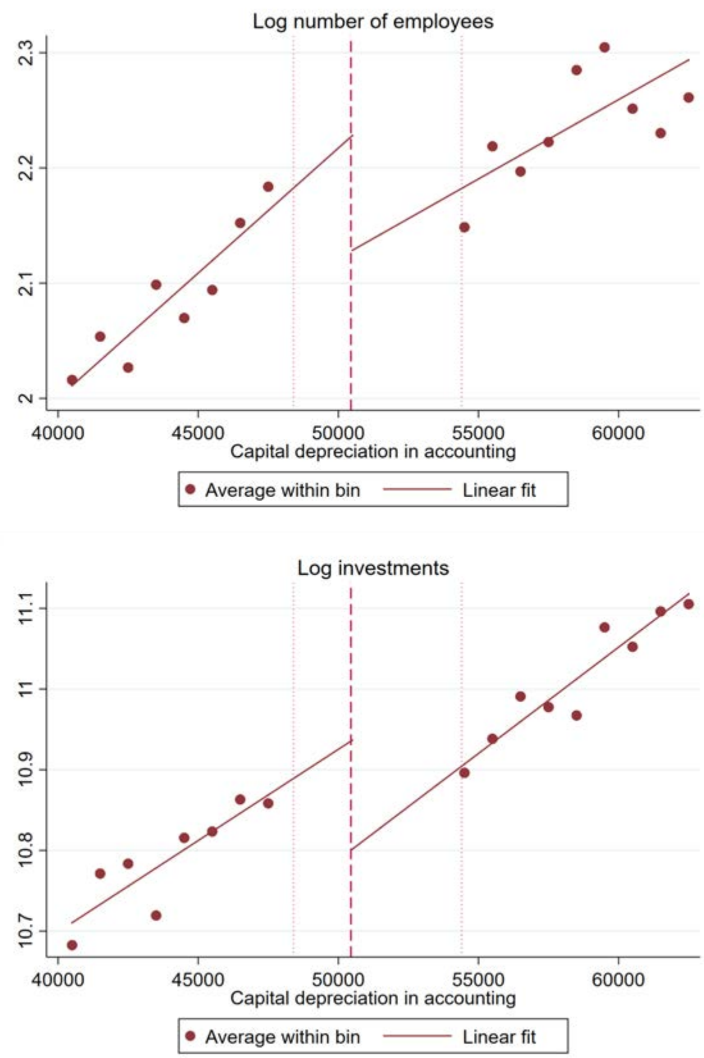
Notes: These Figures plot the estimated response of investment (panel a) and employment (panel b) by size of bandwidth.

Figure 14: Real Versus Reporting Response



Notes: These Figures compare tax depreciation to accounting depreciation measures. The first panel plots the distribution of the difference between tax and accounting depreciation for firms that bunch at the threshold and firms that do not. The second and third panel show the distribution of tax and accounting depreciation, respectively, in the neighborhood of the payroll tax discontinuity.

Figure 15: Accounting depreciations as a running variable



Notes: These Figures plot our main outcomes, the (log) number of employees (first panel) and total investments (second panel), using capital depreciations in accounting as a running variable.

Table 8: Earnings Responses by Employee Types

Outcomes: Mean employee-level log earnings				
Unionization	All	Non-union	Fake union	Union
RD Estimate	-0.008	-0.025*	0.003	-0.001
	(0.014)	(0.015)	(0.014)	(0.013)
Bandwidth	30,000	30,000	30,000	30,000
N above	35,404	28,045	30,526	32,442
N below	67,317	50,133	54,015	58,536
Education	No High School	High School	No College	College degree
RD Estimate	-0.013	0.009	-0.011	-0.002
	(0.013)	(0.016)	(0.013)	(0.014)
Bandwidth	30,000	30,000	30,000	30,000
N above	34,592	24,220	32,168	34,363
N below	65,129	42,185	58,154	64,090
Tasks	Upper level	Lower level	Manual	
RD Estimate	0.027	0.003	-0.016	
	(0.017)	(0.015)	(0.013)	
Bandwidth	30,000	30,000	30,000	
N above	22,006	25,323	29,541	
N below	35,987	44,171	52,013	
Gender	Men	Women		
RD Estimate	-0.049	0.005		
	(0.043)	(0.023)		
Bandwidth	30,000	30,000		
N above	4,985	11,914		
N below	9,708	21,564		

Notes: This Table reports the results of estimating equation (1) on mean employee-level log earnings for all workers ("All"), non-unionized workers ("Non-union"), fake union workers ("Fake union"), unionized workers ("Union"), workers with no high school degree ("No high school"), workers with no college degree ("No College"), workers with a college degree ("College Degree"), upper-level workers ("Upper Level"), lower-level workers ("Lower Level"), manual workers ("Manual"), male workers ("Men") and female workers ("Women"). We use a fixed 30,000 euro bandwidth in these specifications to have comparable estimates across employee types due to the relatively small number of observations in some of the categories.

Table 9: Earnings Responses by Earnings Decile

	Outcomes: Mean employee-level log earnings				
Decile	Smallest decile	2nd	3rd	4th	5th
RD Estimate	-0.128**	-0.021	-0.001	-0.007**	0.001
	(0.064)	(0.032)	(0.005)	(0.004)	(0.003)
Bandwidth	30,000	30,000	30,000	30,000	30,000
N above	1,818	2,715	3,422	3,971	4,139
N below	5,750	6,949	7,391	7,624	7,831
Decile	6th	7th	8th	9th	Largest decile
RD Estimate	-0.002	0.002	-0.007*	0.002	0.018
	(0.003)	(0.003)	(0.004)	(0.010)	(0.023)
Bandwidth	30,000	30,000	30,000	30,000	30,000
N above	4,327	4,093	4,180	3,463	3,276
N below	7,413	6,925	6,818	5,363	5,253

Notes: This Table reports the results of estimating equation (1) on mean employee-level log earnings by deciles of earnings. We use a fixed 30,000 euro bandwidth in these specifications to have comparable estimates across earnings deciles due to the relatively small number of observations in each category.

Table 10: Estimates by Size of the Missing-Mass Region

	Size of the donut hole					
	500		1500		2500	
Outcome	Log investments	Log no. empl.	Log investments	Log no. empl.	Log investments	Log no. empl.
RD Estimate	-0.220***	-0.068*	-0.208***	-0.067	-0.227***	-0.131***
	(0.050)	(0.038)	(0.060)	(0.045)	(0.064)	(0.047)
Bandwidth	7,321	6,297	7,921	6,925	9,082	8,388
N above	12,358	11,039	11,439	10,206	11,391	10,717
N below	18,469	15,966	20,326	17,982	24,057	22,697
	Size of the donut hole					
	3500		4500			
Outcome	Log investments	Log no. empl.	Log investments	Log no. empl.		
RD Estimate	-0.178***	-0.122***	-0.147***	-0.049***		
	(0.047)	(0.034)	(0.034)	(0.019)		
Bandwidth	16,620	14,011	29,879	42,114		
N above	20,560	17,615	34,154	46,061		
N below	52,899	43,635	104,006	108,741		

Notes: This Table reports the results of estimating equation (1) on employment and investment using different thresholds for the missing-mass region (donut hole to the right of the threshold). The Table shows bias-corrected estimates with robust standard errors, the size of the optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within the bandwidth, respectively, following Calonico et al. [2014].

Table 11: Estimates by Size of the Excess-Mass Region

	Size of the excess-mass region							
	3500		3000		2500		2000	
Outcome	Log investments	Log no. empl.	Log investments	Log no. empl.	Log investments	Log no. empl.	Log investments	Log no. empl.
RD Estimate	-0.174***	-0.065**	-0.136***	-0.063**	-0.114**	-0.066**	-0.110**	-0.079***
	(0.059)	(0.033)	(0.053)	(0.032)	(0.048)	(0.029)	(0.044)	(0.030)
Bandwidth	16,074	18,388	17,868	18,524	19,262	19,604	20,501	18,442
N above	19,289	23,309	21,820	23,492	23,603	24,856	25,149	23,376
N below	43,211	56,486	52,853	58,682	61,423	66,060	69,592	61,187

	Size of the excess-mass region					
	1500		1000		500	
Outcome	Log investments	Log no. empl.	Log investments	Log no. empl.	Log investments	Log no. empl.
RD Estimate	-0.115***	-0.081***	-0.139***	-0.089***	-0.147***	-0.123***
	(0.042)	(0.028)	(0.043)	(0.028)	(0.041)	(0.031)
Bandwidth	21,066	19,019	20,376	19,011	21,062	16,575
N above	25,819	24,125	24,993	24,115	25,814	20,687
N below	74,234	65,811	71,862	67,368	77,555	56,931

Notes: This Table reports the results of estimating equation (1) on employment and investment using different thresholds for the excess-mass region (donut hole to the left of the threshold). The Table shows bias-corrected estimates with robust standard errors, the size of the optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within the bandwidth, respectively, following Calonico et al. [2014].

Table 12: Estimates by Different Polynomials of Different Degrees

	Polynomial fit			
	2nd degree		3rd degree	
Outcome	Log investments	Log no. empl.	Log investments	Log no. empl.
RD Estimate	-0.128***	-0.086***	-0.212***	-0.115***
	(0.040)	(0.026)	(0.048)	(0.036)
Bandwidth	20,362	19,578	38,541	36,819
N above	24,978	24,828	43,151	43,070
N below	71,791	70,433	104,006	108,741

Notes: This Table reports the results of estimating equation (1) on employment and investment using different polynomial fits. The Table shows bias-corrected estimates with robust standard errors, the size of the optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within the bandwidth, respectively, following Calonico et al. [2014].

Table 13: Capital Depreciations in Accounting as a Running Variable

	Capital depreciations in accounting	
Outcome	Log investments	Log no. empl.
RD Estimate	-0.121***	-0.099***
	(0.031)	(0.048)
Bandwidth	35,445	35,039
N above	27,263	26,581
N below	54,533	52,123

Notes: This Table reports the results of estimating equation (1) on employment and investment using capital depreciations in accounting as a running variable. The Table shows bias-corrected estimates with robust standard errors, the size of the optimal bandwidth (one common mean square error optimal bandwidth), the number of observations above (N above) and below (N below) the threshold within the bandwidth, respectively, following Calonico et al. [2014].

Table 14: Social insurance percentages by firm categories, different insurance types and years

	Health and pension				Unemployment					
	Firm categories*			Accident	Firm categories**		Group life	Employees	Total	Total
Year	I	II	III	insurance***	I	II	insurance***	pension***	lowest	highest
1996	4.000	5.600	6.500	1.2	1.00	4.00	0.100	16.80	23.100	28.600
1997	4.000	5.600	6.500	1.4	1.00	4.00	0.090	16.70	23.190	28.690
1998	4.000	5.600	6.500	1.4	0.90	3.90	0.080	16.80	23.180	28.680
1999	4.000	5.600	6.500	1.3	0.90	3.85	0.080	16.80	23.080	28.530
2000	4.000	5.600	6.500	1.2	0.90	3.45	0.090	16.80	22.990	28.040
7/2000	3.600	5.600	6.500	1.2	0.90	3.45	0.090	16.80	22.590	28.040
2001	3.600	5.600	6.500	1.2	0.80	3.10	0.095	16.60	22.295	27.495
2002	3.600	5.600	6.500	1.1	0.70	2.70	0.095	16.70	22.185	27.085
3/2002	2.950	5.150	6.050	1.1	0.70	2.70	0.095	16.70	21.535	26.635
2003	2.964	5.164	6.064	1.1	0.60	2.45	0.081	16.80	21.545	26.495
2004	2.964	5.164	6.064	1.1	0.60	2.50	0.080	16.80	21.544	26.544
2005	2.966	5.166	6.066	1.2	0.70	2.80	0.080	16.80	21.746	26.946
2006	2.958	5.158	6.058	1.1	0.75	2.95	0.080	16.70	21.588	26.888
2007	2.951	5.151	6.051	1.1	0.75	2.95	0.080	16.64	21.521	26.821
2008	2.771	4.971	5.871	1.0	0.70	2.90	0.080	16.80	21.351	26.651
2009	2.801	5.001	5.901	1.0	0.65	2.70	0.070	16.80	21.321	26.471
4/2009	2.000	4.201	5.101	1.0	0.65	2.70	0.070	16.80	20.520	25.601
2010	2.220	2.220	2.220	0.8	0.75	2.95	0.070	16.90	20.74	22.94
2011	2.210	2.210	2.210	1.0	0.80	3.20	0.070	17.10	21.18	23.58
2012	2.210	2.210	2.210	1.0	0.80	3.20	0.070	17.35	21.43	23.83
2013	2.040	2.040	2.040	0.9	0.80	3.15	0.070	17.35	21.16	23.51
2014	2.140	2.140	2.140	0.9	0.75	2.95	0.070	17.75	21.61	23.81
2015	2.080	2.080	2.080	0.9	0.80	3.15	0.070	18.00	21.85	24.89
2016	2.120	2.120	2.120	0.8	1.0	3.90	0.070	18.00	21.99	24.89
2017	1.080	1.080	1.080	0.8	0.8	3.30	0.070	17.95	20.70	23.20

* Refers to firm categories by wage sums and capital depreciation.
** Category I is for wages below a certain wage sum threshold, e.g. 2,059,500 euro in year 2017, and Category II is for wages above the threshold. The threshold varies slightly over years.
*** Represents the average values of these insurances.

Table 15: Descriptive statistics: Firm-level sample vs. all Finnish firms in 2002

Sample							
VARIABLE Statistics	Depreciations in taxation	Depreciations in accounting	Capital Stock	Investments	Investments Fixed assets	Investments Buildings	Investments R&D
Mean	49627.8	41413.3	262603.9	90755.2	66497.8	17836.2	6421.2
Median	46634.5	44876.1	176874.5	57025.7	43283.8	0	0
Se. mean	135.4	560.5	8136.1	3364.3	1673.2	2895.1	600.5
	Sales	Intermediate costs	Labor costs	Number of employees	Profits	Value added	Labor Productivity
Mean	1132873	657483	205057.1	12.4	52490.7	475440	1.330
Median	561326.5	150501	156830.3	9	30872.7	379644.9	1.138
Se. mean	38766.3	36010.5	3118.6	.235	4301.0	7738.6	.017
N=2,972							

All Finnish firms							
VARIABLE Statistics	Depreciations in taxation	Depreciations in accounting	Capital Stock	Investments	Investments Fixed assets	Investments Buildings	Investments R&D
Mean	31249.6	27384.5	228694.9	58058.6	31461.2	18564.5	8032.8
Median	436.7	0	0	0	0	0	0
Se. mean	891.1	1008.3	13265.1	6080.1	1471.5	5624.6	951.2
	Sales	Intermediate costs	Labor costs	Number of employees	Profits	Value Added	Labor Productivity
Mean	606997.3	306996.0	102499.5	10.1	42562.6	300001.4	.968
Median	66537.0	3355.4	3124.8	1	3105.9	47954.8	.650
Se. mean	9730.6	6330.7	1841.2	.360	5009.9	5000.5	.003
N=148,211							

Notes: The upper panel of this Table reports the descriptives statistics for the data used in the graphical analysis in the paper. The sample is restricted to firms with capital depreciations between 40,500–64,500 euros, and excluding the donut hole region. The lower panel of the Table shows the same descriptive statistics for all Finnish firms with sales between 10,000–100,000,000 euros. The descriptive statistics are presented only for year 2002, the mid-year of our treatment period 1996–2009.

Table 16: Descriptive statistics: Industry and organizational form distribution

	Below the threshold			Above the threshold		
Industry classification	Frequency	Share	Cumulative	Frequency	Share	Cumulative
Farming & Mining	2,389	8.93	8.93	1,920	9.91	9.91
Manufacturing	4,462	16.68	25.61	3,454	17.83	27.75
Construction & Transportation	14,356	53.66	79.27	9,915	51.19	78.94
Services	3,402	12.72	91.99	2,427	12.53	91.47
Finance & Real estate	1,883	7.04	99.02	1,472	7.60	99.07
Other & Missing	261	0.98	100.00	181	0.93	100.00
Organizational form	Frequency	Share	Cumulative	Frequency	Share	Cumulative
Sole proprietors	2,933	10.96	10.96	1,682	8.68	8.68
Corporations	19,185	71.68	82.64	14,613	75.43	84.11
Partnerships	4,647	17.36	100.00	3,079	15.89	100.00

Notes: This Table reports the number of firms, the share of firms and cumulative proportion of firms by industry codes and organizational form for the data used in the graphical analysis in the paper. The sample is restricted only to firms with capital depreciations between 40,500–64,500 euros, and excluding the donut hole region.

Table 17: Definitions of the variables used in the analysis

Variables	Definitions
Payroll tax rate	Firm-level payroll tax rate for health and pension contributions.
Capital depreciation in taxation	Firm-level annual capital depreciations used in taxation in euros.
Capital depreciation in accounting	Firm-level annual capital depreciations in accounting in euros.
Earnings	Employee-level total annual earnings of employees.
Labor costs	Annual total wages and other wage-related compensations paid by the firm to employees excluding all social insurance contributions and taxes in euros.
Number of employees	The sum of the number of employees who worked in the firm during the tax year.
Secondary degree	Employee-level education measure for individuals with a bachelor or masters degree or higher.
High school graduate	Employee-level education measure for individuals who have graduated from high school.
Upper-level employees	Employee-level task measure for individuals whose position is senior official and upper management, senior officials and employees in research and planning, senior officials and employees in education and training or other senior officials and employees.
Lower-level employees	Employee-level task measure for individuals whose position is supervisor, clerical and sales workers or independent work.
Manual workers	Employee-level task measure for individuals whose position is clerical and sales worker, worker in agriculture, forestry and commercial fishing, manufacturing worker, other production worker or distribution and service worker.
Investments	Annual euro value of gross investments in fixed capital, buildings and research and development.
Fixed asset investments	Annual euro value of gross investments in machines and equipment.
Building investments	Annual euro value of gross investments in buildings.
R&D investments	Annual euro value of gross investments in research and development.
Sales	Gross annual sales of the firm from its primary operating activity minus any discounts given, valued-added taxes, and other taxes based on sales volumes.
Intermediate inputs	Annual euro value of the costs used as intermediate inputs in production.
Labor productivity	Annual euro value of sales minus intermediate inputs divided by labor costs.
Capital productivity	Annual euro value of sales minus intermediate inputs divided by annual investments.
Union employees	Employee-level dummy for individuals with above-median tax deductible labor union membership fee.
Fake union employees	Employee-level dummy for individuals with below-median tax deductible labor union membership fee.
Not union employees	Employee-level dummy for individuals with no tax-deductible labor union membership fee.