Are Taxes Turning Humans Into Machines? Using Payroll Tax Variation to Estimate the Capital-Labor Elasticity of Substitution.*

Youssef Benzarti and Jarkko Harju†

June 26, 2018

Abstract

This paper provides a quasi-experimental estimate of the micro capital-labor elasticity of substitution. We use a discontinuity in the average payroll tax rate to show that firms decrease both labor and capital when experiencing higher payroll tax rates. We argue that this is consistent with Leontief production functions, implying a micro capital-labor elasticity of substitution of zero. Using a conceptual framework to aggregate micro elasticities, we estimate a macro capital-labor elasticity of substitution of 0.17. Our finding sheds new light on the role of the substitution of labor with capital in the fall of labor shares and on the fiscal externalities imposed by payroll taxes.

*We are grateful to Stefanie Stantcheva and Alisa Tazhitdinova for helpful discussions.
†Benzarti: UCLA and NBER, benzarti@econ.ucla.edu; Harju: VATT Institute for Economic Research (Helsinki, Finland) and CESifo, jarkko.harju@vatt.fi;
1 Introduction

There is a long tradition in economics of estimating production functions, dating back to at least the early 1800’s. More recently, there has been renewed interest in the question and, in particular, in estimating the elasticity of substitution between capital and labor. Part of this interest stems from explaining a puzzling stylized fact: the steady decline in the labor share of income. For example, Piketty [2014] and Piketty and Zucman [2014] argue that falling labor shares can be explained by an elasticity of substitution that is greater than one, which allows firms to substitute one unit of labor with less than one unit of capital and therefore could explain the declining labor share patterns. This elasticity of substitution is also important from a policy perspective as it governs how employment and investment respond to any policy shocks that affect the price of labor or capital, such as payroll tax changes, minimum wage laws or depreciation rules.

It is notoriously difficult to estimate the elasticity of substitution between capital and labor. One approach has been to estimate it using aggregate time series with structural assumptions on the production function. This approach yields a wide range of elasticities that essentially depend on the assumptions made. A second approach has been to use micro data to estimate production functions. However, papers using this approach often assume Cobb-Douglas production functions in their estimation, essentially setting the elasticity of substitution to 1. The remaining papers suffer from the difficulty of finding plausibly exogenous variation in factor prices.\footnote{1}

In this paper, we use quasi-experimental variation in labor costs to estimate the micro capital-labor elasticity of substitution.\footnote{2} First, we find that labor and capital are complements at the micro level: we show that firms systematically reduce both labor and capital when payroll

\footnote{1}{For example Raval [2014] and Oberfield and Raval [2014] estimate the capital-labor elasticity of substitution using micro data by relying on cross sectional variation in local wages. While this approach substantially improves upon the previous literature in terms of identification, it is still likely to be biased as location choices of firms are presumably endogenous to labor costs.}

\footnote{2}{In ongoing, independent and contemporaneous work, Moreau [2018] is using the same variation as in Garicano et al. [2016] to estimate the capital-labor elasticity of substitution in France.}
taxes increase exogenously. This is consistent with Leontief production functions, i.e., CES production functions with a capital-labor elasticity of substitution equal to zero. This stands in contrast with previous estimates as the literature usually assumes or estimates positive elasticities of substitution. Second, we uncover some degree of heterogeneity in the response of labor and capital to changes in labor costs: (1) there is a positive correlation between the number of employees and the magnitude of the response of capital to payroll taxes; (2) we find an inverse-U shape relationship between firm productivity – as measured by value-added, turnover or profits per employee – and the magnitude of the response of capital flows to payroll taxes. Third, using the framework from Oberfield and Raval [2014] we aggregate our micro estimates to derive a macro estimate of the capital-labor elasticity of substitution, that would account for capital-labor substitution across firms. We estimate an aggregate substitution elasticity of 0.17.

The findings of this paper are important for the following reasons. First, there are no quasi-experimental estimates of the capital-labor elasticity of substitution despite of this elasticity being a key academic and policy parameter. Harberger [1962], for example, shows that this elasticity is a key determinant of the incidence of corporate taxes in an open economy. Second, our empirical estimate of this elasticity help settle the debate over whether the falling labor shares of income can be rationalized with an elasticity of substitution greater than one. While the literature has estimated both elasticities that are greater than one (Karabarbounis and Neiman [2013]) and smaller than one (Lawrence 2015; Oberfield and Raval [2014]; Antràs 2004; Hamermesh 1996), our paper shows that the micro elasticity of substitution is zero and the macro elasticity very close to zero. Our finding thus casts doubt on capital-labor substitution being the driver of the falling labor share. Third, our paper contributes to a Public Finance literature concerned with the efficiency of payroll taxes. While previous research has estimated the distortionary effects of payroll taxes on wages and employment (Saez et al. 2012 and Saez et al. 2017), few have estimated the fiscal externality imposed by payroll taxes. This paper shows that payroll taxes impose a positive fiscal externality: as payroll taxes increase, firms reduce their flow of capital therefore decreasing their depreciation which results in an increase in revenue from corporate taxes. Accounting for this effect is important when considering payroll tax reforms. More generally,
our estimates are important for government estimates of projected shortfalls for government programs that are funded by payroll taxes.

2 Conceptual Framework

2.1 Micro Capital-Labor Elasticity of Substitution

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

\[ F(k, l) = \left( \alpha k^{\frac{\sigma-1}{\sigma}} + (1 - \alpha)l^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \]

where \( k \) is capital, \( l \) is labor, and \( \alpha \) and \( \sigma \) are parameters. \( \sigma \) is assumed to be strictly positive and has no upper bound. When \( \sigma \rightarrow 0 \), 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 \( \text{RTS} \) the rate of technical substitution between capital and labor. It can be shown that the capital-labor substitution elasticity only depends on \( \sigma \):

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

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 wage and \( r \) is the cost of capital:
\[
\min_{k,l} C(w, r) = w l + r k
\]
subject to
\[
F(k, l) = q_0
\]
This yields the following condition:
\[
k = \left( \frac{w \alpha}{r (1 - \alpha)} \right) \sigma \ l
\]
Using this relationship and the resource constraint \( F(k, l) = q_0 \), we get:
\[
l = q_0 \left( \alpha \left( \frac{w \alpha}{r (1 - \alpha)} \right)^{\sigma - 1} + (1 - \alpha) \right)^{\frac{\sigma}{1 - \sigma}}
\]
\[
k = q_0 \left( (1 - \alpha) \left( \frac{w \alpha}{r (1 - \alpha)} \right)^{1 - \sigma} + \alpha \right)^{\frac{\sigma}{1 - \sigma}}
\]
We take the derivative of these two equations with respect to \( w \) to get the elasticity of capital and labor with respect to wage:
\[
\epsilon_{k,w} = \frac{\partial k}{\partial w} \frac{w}{k} = \frac{(1 - \alpha) \sigma}{(1 - \alpha) + \alpha \left( \frac{w \alpha}{r (1 - \alpha)} \right)^{\sigma - 1}}
\]
\[
\epsilon_{l,w} = \frac{\partial l}{\partial w} \frac{w}{l} = -\frac{\alpha \sigma}{\alpha + (1 - \alpha) \left( \frac{w \alpha}{r (1 - \alpha)} \right)^{1 - \sigma}}.
\]
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 productions 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 increases when labor costs increase, then $\epsilon_{k,l} > 0$. If instead, investments decreases when labor costs increase then $\epsilon_{k,l} = 0$. In the rest of the paper, we setup our empirical framework to estimate how investments respond to changes in labor costs.

3 Data and Institutional Background

3.1 Institutional Background

In Finland, both employees and employers contribute to social insurance. Social insurance contributions include contributions to pension schemes, unemployment insurance, accident insurance, health insurance and life insurance. The employer portion of social insurance contributions are based on the annual salaries paid to their employees. In general, the employer’s statutory share of total contributions is larger than that of their employees. Total contributions are split between employees and employers, but the split depends on several firm and worker characteristics, including, for example, the age of the worker. For example, in 2017, the average pension insurance contribution rate was 17.95 percent of a

\footnote{The largest share of total social insurance contributions goes to pension contributions.}
given employee’s monthly gross wage and the employee’s contribution rate is 7.65 percent when older than 53 and younger than 62, otherwise the employee’s contribution rate is 6.15%. Appendix Table 3 shows the employers’ social insurance contribution rates by firm categories, for different insurance types and years.

Before 2010, the payroll contribution rate for national health and pension of employers depended on total labor costs and the the level of capital depreciation, as shown in Table I below, while workers’ tax rates were not affected by these discontinuities. Category I corresponds to firms with less than 50,500 euros of annual capital depreciation (D) or more than 50,500 euros but less than 10% of annual salaries. Category II corresponds to firms with depreciation levels of more than 50,500 euros and 10 to 30% of labor costs. When the depreciation level exceeds 50,500 euros 30% of labor costs, social security contributions are paid according to the third category. Appendix Table 3 shows, for example, that the contribution rates to the national health and pension schemes by firm categories were 2.964%, 5.164% and 6.064% in 2004, respectively.

In 2010, these categories were removed when the government decided to remove the national pension insurance contributions and the rate was set to be 2.23% for all firms irrespective of capital depreciation and salaries.

<table>
<thead>
<tr>
<th>Definition for firm categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
</tr>
<tr>
<td>II</td>
</tr>
<tr>
<td>III</td>
</tr>
</tbody>
</table>

Table 1: Firm categories for payroll tax rates
Note: D refers to tax deductible capital depreciations and labor costs refer to all salaries.

There are two important details. First, these categories were determined by the latest available tax information and salaries paid for the same year as the one used to determine depreciation levels. Therefore, for example, the 2006 payment category is based on fiscal year 2004. New firms were always subject to the category I for the first two years of operation, irrespective of their depreciation levels and labor costs.

Second, what counts as depreciation for tax purposes can be different from what counts

\footnote{We provide details of the depreciation rules in Appendix Section A}

\footnote{See legislation in Finlex: HE 147/2009}
as depreciation for accounting purposes. Depreciation in accounting is a systematic reduction of the cost of a fixed asset. According to the Finnish tax law, the amount of annual depreciation for tax purposes cannot be larger than that for accounting purposes. This opens up a possibility for firms to manipulate the amount of annual taxable depreciation, e.g. for tax planning purposes to avoid higher payroll taxes. Depreciation for accounting purposes, however, cannot be manipulated and is subject to strict auditing requirements. Fortunately, we have data on both of these variables and we can examine the extent to which this manipulation exists. We do this evaluation in Section 4.

3.2 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 some firm characteristics including organizational form, location and industry code. The data provides yearly information, at the firm level, on labor costs, number of employees, both book (accounting) and tax amounts of capital depreciation and the level of capital investment. In addition, we have firm level data on sales and various cost categories, including material and rental costs as well basic firm characteristics, such as industry code and firm location.

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% of all wages. Legally, the discontinuity we consider does not apply to these firms so there is no reason to include them. This restriction removes approximately 25% of the total data.

4 Results

4.1 Empirical Approach

In Section 2, we showed that the key parameter of interest is how capital investments respond to changes in labor costs. If investments increase when labor costs increase, firms are
substituting workers with capital, implying a positive capital-labor elasticity of substitution. If instead, capital investments decrease when labor costs decrease, firms are operating with fixed proportions of capital and labor, which corresponds to Leontief production functions. To estimate the response of capital investment with respect to labor costs, we use the discontinuity in payroll taxes due at the €50,500 depreciation threshold as described in Section 3.1. As firms cross the €50,500 depreciation threshold, the marginal and average payroll tax rates discontinuously increase, effectively increasing labor costs. While in principle firms can adjust wages down by the full amount of the payroll tax increase, our estimates on labor costs show that the pass-through of payroll taxes to wages is certainly far from 100% and likely close to zero.⁶ We also estimate a negative effect on employment, confirming the existence of a “first stage”.

Because our running variable (depreciation) can be manipulated by firms, 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 methods from Kleven and Waseem [2013] to identify the manipulated area which, in their framework, corresponds to the area of the excess and missing masses. We describe this approach in detail in Section 5.2.2.

We follow the approach of Calonico et al. [2014] to estimate the optimal bandwidth and report bias-corrected with robust standard error estimates. In addition, we perform placebo tests by running our specification on post-2010 years when the payroll tax discontinuity was removed. Formally, we run the following regression separately on the 1996 to 2009 period when the discontinuity was in place and on the post 2010 period, when the discontinuity was removed:

\[
\log(y_i) = \alpha + \beta_1 \cdot (\text{depreciation}_i - d) + \beta_2 \cdot \text{Above}_i + \beta_3 \cdot \text{Above}_i \cdot (\text{depreciation}_i - d) + \epsilon_i \quad (1)
\]

⁶We estimate the incidence of payroll taxes using this discontinuity and another variation in Benzarti and Harju (2018b) and find limited pass-through.
where \( y_{it} \) is the outcome of interest for firm \( i \), \textit{depreciation} is the level of capital depreciations, \( d \) is the depreciation threshold above which the average payroll tax rate increases, \textit{Above} is a dummy (1 above the depreciation threshold, 0 otherwise), \( \epsilon_i \) is error term and is calculated using the robust standard errors from Calonico et al. [2014]. \( \beta_3 \) is the coefficient of interest showing the magnitude of the change of the outcome variable at the discontinuity.

### 4.2 Graphical Evidence

Figure 1 plots the average payroll tax above and below the €50,500 depreciation threshold. Average payroll taxes increase at the threshold, confirming the existence of the discontinuity we need in our RD design.

Figures 2, 3 and 2 show the effect of the increase in payroll taxes on our main outcomes of interest: number of employees, flow of capital (investment) and turnover, respectively. The estimation of the red fitted lines excludes firms that are located in the bunching area, consistent with the donut hole RD approach. The first panel pools years 1996 to 2009, when the payroll tax discontinuity was in place, while the second panel pools post-2010 years, when there was no discontinuity in payroll taxes, and is used as a placebo test.

The first panel of Figure 2 shows a decrease in number of employees as payroll taxes increase implying that \( \epsilon_{t,w} \) is negative. The second panel of Figure 2 shows that there is no discontinuity in number of employees during the placebo period. This helps mitigate concerns that are our results are spurious.

Figure 3 shows the effect of the increase in average payroll tax rates on investment for both the treatment and placebo periods. We observe a discontinuous decrease in investments at the depreciation threshold for the 1996-2009 period and we visually detect no decrease in investments during the placebo period. This observed decrease in capital flows implies that firms are reducing their capital when labor costs increase, which is consistent with Leontief production functions where capital and labor are used in fixed proportions as discussed in Section 2.

Finally, we also analyze the effect of higher labor costs on sales. Figure 4 shows the response of sales above and below the depreciation threshold. There is a decrease in sales in the 1996-
2009 period relative to the post-2010 placebo period, suggesting that, as firms decrease their number of employees and investment because of the higher labor costs, they also reduce their output. However, there are two caveats to interpreting this finding as evidence of firms operating with Leontief production functions. First, the decrease in sales could be consistent with both Leontief production functions and CES production functions with a positive capital-labor elasticity that is smaller than 1. If capital is not a perfect substitute for labor, an increase in labor costs would also lead to a decrease in output. Second, this interpretation relies on the assumption that payroll taxes are not passed through to price i.e. that prices are not different above and below the depreciation threshold. In principle, an increase in payroll taxes could lead to an increase in prices and a decrease in quantities, which could result in either an increase or a decrease in sales.

4.3 Regression Estimates

Table 2 shows the result of running specification on the treatment years (1996 to 2009) and the placebo years (post 2010) separately and is the regression counterpart of the graphical evidence discussed in Section 4.2. We report conventional, bias-corrected and robust bias-corrected estimates following the approach of Calonico et al. [2014]. The estimates for the treatment period are systematically negative and clearly statistically significant, while the placebo estimates are substantially smaller in magnitude and statistically insignificant. The effect on sales is robust to the bias correction of Calonico et al. [2014] but their approach of calculating standard errors doubles them making the coefficients insignificant. Nevertheless, the estimated effect on turnover for the treatment years is double that of the placebo years suggesting that payroll taxes have some effect on turnover.

Our estimates imply that the average payroll tax rate increase causes firms to reduce their number of employees by 8.38% and their investment by 14.5%. This is consistent with the graphical evidence discussed in Section 2. The fact that both labor and capital are reduced when the cost of labor increases is consistent with Leontief production functions and inconsistent with positive capital-labor elasticities of substitution as discussed in Section 2. The effect on turnover, while not robust to different specifications, is large, as firms reduce
their output by 4.9% on average.

5 Heterogeneity

To estimate the extent to which there is heterogeneity in the response of firms to the increase in average payroll tax rates we rely on two approaches. First, we perform the empirical analysis outlined in Section 4.1 on quartiles of firms. We use several proxies for firm size to test for whether larger firms are more likely to exhibit a positive capital-labor elasticity of substitution. Second, we use the degree of manipulation at the threshold and a bunching estimator. While this approach requires additional assumption to estimate the absolute magnitude of the capital-labor elasticity of substitution, it has the advantage of requiring less power than RD designs and is informative of the relative magnitude of the response. It allows us to further break down firms into smaller quantiles and better assess the degree of heterogeneity.

5.1 RD Heterogeneity

We use four different measures of firm size: sales, profits, markups and labor costs. Sales, profits and labor costs are taken directly from the corporate tax data. Sales are defined as price multiplied by quantity, profits are gross of taxes and are equal to sales minus all deductible costs, and labor costs are inclusive of the employer and employee portion of payroll taxes and are equal to the total expenses employers spend on workers. We define markups as the difference between turnover and variables costs divided by variable costs. We run specification 1 on subsamples of our main dataset by quartiles of these four variables. Our main parameter of interest is the sign of the response of capital flows to the increase in labor costs. A negative coefficient implies that production functions are Leontief, while a positive one would imply a positive capital-labor elasticity of substitution.

We summarize our findings in Figure 5, which plots the magnitude of the response of capital flow for each quartile of our different measures of firm size. 

Regression outputs are readily available upon request.
in the magnitude of the response of capital flows to the increase in labor costs, the point estimates are systematically negative for all the subsamples we consider and we detect no specific size patterns. Depending on the subsamples, we can either reject that the response is weakly positive or fail to reject that it is zero with an upper bound on the 95% confidence intervals of at most 0.2. Hence, it is unlikely that the capital-labor elasticity of substitution is positive for larger or smaller firms, conditional on the assumption that markups, sales, wage bills and profits are good proxies for size. Given that we only have four groups, it is harder to derive any specific relationships between the response of capital to labor costs and firm sizes. Next, we turn to a bunching estimator which allows us to break down our samples even further.

5.2 Bunching and Elasticity

5.2.1 Conceptual Framework

The interpretation of the magnitude of bunching depends on whether the capital-labor elasticity of substitution is positive or zero. If it is positive, then we show in Appendix Section 3 that the magnitude of bunching and the capital-labor elasticity of substitution are positively correlated. Intuitively, it is easier to substitute away from labor towards capital when the elasticity of substitution is large and hence avoid the higher level of payroll taxes by bunching at the threshold. When the capital-labor elasticity of substitution is zero, the only possible source of variation in bunching is due to heterogeneity in the response of labor to the payroll tax rate, i.e., bunching is positively correlated with $\epsilon_{l,\tau}$. Given that both the evidence from Section 4.1 and Section 5.1 imply Leontief production functions, we interpret our evidence in light of the capital-labor elasticity of substitution being zero. Therefore, any variation in bunching is interpreted as being due to heterogeneity in $\epsilon_{l,\tau}$ rather than heterogeneity in the capital-labor elasticity of substitution.

5.2.2 Bunching Methodology

We follow Chetty et al. [2011] and Kleven and Waseem [2013] to estimate 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 to 100 euros bins and count the number of firms in each bins. Then we estimate a counterfactual density by running the following regression while excluding the region around the threshold \([D_L, D_H]\):  

\[
c_j = \sum_{i=0}^{p} \beta_i (D_j)^i + \sum_{i=D_L}^{D_H} \eta_i \cdot 1(D_j = i) + \varepsilon_j
\]  

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. Intuitively, this is the point at which the behavioral response starts.

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 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 boot-
strapped 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.

5.2.3 Results

Figure 6 shows the bunching estimates by deciles of markup, turnover per employee, value-added per employee and total labor costs. Two main patterns emerge. First, as can be seen in the first panel of Figure 6, the magnitude of bunching increases with the number of employees. This is most likely due to the fact that, as the number of employees increases, the incentives to bunch are larger as firms are paying higher total payroll taxes. Second, once we normalize firm performance measures by the number of employees, namely value added per employee and turnover by employee, we find that the magnitude of bunching follows an inverse U-shape pattern. Median firms are more likely to bunch relative to firms in lower or higher deciles. This pattern could be consistent with less productive firms having relatively large labor adjustment costs and very productive firms having invaluable employees too costly to part with. We find similar inverse U-shaped patterns for markup deciles.

5.2.4 Real or Reporting Response?

In principle, bunching can be due to reporting rather than real responses. However, there are two types of depreciation: depreciation for tax purposes and for accounting purposes. The payroll tax discontinuity is based on the tax depreciation level. In principle, it is possible to manipulate tax depreciation without changing “real” depreciation by electing to use a different tax depreciation schedule (see details in [A]). However, accounting depreciation rules are not flexible and there is very little room for evasion along this margin as firms are subject to strict annual auditing requirements. For this reason, any response in accounting depreciation is likely to be a real response rather than a reporting one.

In our dataset, we have access to both tax and accounting depreciation levels. The first panel of Figure 7 shows the difference between these two variables for bunchers and for non-bunchers. The distributions look very similar, suggesting that the response of bunchers
is likely to be real. Moreover, the second and third panels show the distribution of firms at the tax depreciation threshold using tax depreciation (second panel) and the accounting depreciation (third panel) as running variables. The magnitude of bunching is very similar for the tax and accounting depreciation, suggesting that the response to the threshold is likely to be real.

6 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 and or 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 \( \alpha_i = \frac{r_k}{r_k + w_l} \) and \( \alpha = \frac{r_k}{r_k + w_l} \) the capital share in the total costs of production for firm \( i \) and the aggregate capital share, respectively. Further, we define \( \theta_i \) 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 \( \sigma^{agg} \) is a weighted average of the micro elasticity of substitution and the elasticity of demand \( \varepsilon \):

\[ \sigma^{agg} = (1 - \chi)\sigma + \chi\varepsilon \]

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 \( \varepsilon \) determines the extent to which consumers respond to lower prices by
shifting consumption to the labor intensive commodity.

We estimate $\alpha_i$, $\alpha$ and $\theta_i$ directly from our corporate tax data, which reports both labor and capital costs. To estimate $\varepsilon$, we use the average markup $\mu$ and assume that $\varepsilon = 1/\mu$. We follow Antras et al. [2017] and define markups as $\frac{\text{sales} - \text{costs}}{\text{costs}}$. We estimate that $\chi = 0.13$ and $\varepsilon = 1.29$. These estimates imply a macro capital-labor elasticity of substitution $\sigma^{agg} = 0.17$.

7 Conclusion

In this paper, we use an exogenous increase in the average payroll tax rate faced by firms to estimate the capital-labor elasticity of substitution. We find that, as payroll taxes increase, both labor and investments decrease, which is consistent with a CES production function with a zero capital-labor elasticity of substitution i.e. a Leontief production function. We also aggregate our micro elasticity and find macro elasticities ranging from 0.1 to 0.3.

Our finding has several implications. First, our finding rules out the possibility that substitution between capital and labor could explain the observed fall in labor shares. Instead, our small substitution elasticity estimates suggest that other mechanisms could be at force. Second, 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 sales and profits. This effect should be taken into account when the Government scores payroll tax changes.
References


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. 2011.


Robert Z Lawrence. Recent declines in labor’s share in us income: a preliminary neoclassical account. 2015.


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


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


Notes: This Figure plots the average payroll tax rates above and below the capital depreciation threshold.
Notes: The first panel shows the response of the number of employees (in log) at the payroll tax discontinuity from 1996 to 2009. The second panel shows a placebo test for years 2010 to 2015 for the same variable.
Figure 3: Investments by bins of capital depreciation.

Notes: The first panel shows the response of the investments (in log) at the payroll tax discontinuity from 1996 to 2009. The second panel shows a placebo test for years 2010 to 2015 for the same variable.
Notes: The first panel shows the response of the sales (in log) at the payroll tax discontinuity from 1996 to 2009. The second panel shows a placebo test for years 2010 to 2015 for the same variable.
Figure 5: Investment Response to Payroll Tax by Measures of Firm Size

Notes: These Figures plot the estimated response of investment to the payroll tax discontinuity using quartiles of labor costs (first panel), turnover (second panel), profits (third panel) and markups (fourth panel).
Figure 6: Bunching Estimates by Measures of Firm Size

Notes: These Figures plot the magnitude of bunching at the depreciation threshold by deciles of labor costs (first panel), markup (second panel), turnover per employee (third panel) and value added per employee (fourth panel).
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 shows the distribution of tax and accounting depreciation, respectively, in the neighborhood of the payroll tax discontinuity.
Table 2: RD donut hole estimates

<table>
<thead>
<tr>
<th>Variables</th>
<th>Treatment (1996 to 2009)</th>
<th>Placebo (post 2010)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>(in logs)</td>
<td>No. employees Investments Turnover</td>
<td>No. employees Investments Turnover</td>
</tr>
<tr>
<td>Conventional</td>
<td>-0.0378*** (0.0140)</td>
<td>-0.0920*** (0.0123)</td>
</tr>
<tr>
<td>Bias-corrected</td>
<td>-0.0838*** (0.0140)</td>
<td>-0.145*** (0.0123)</td>
</tr>
<tr>
<td>Robust</td>
<td>-0.0838*** (0.0272)</td>
<td>-0.145*** (0.0413)</td>
</tr>
<tr>
<td>Observations</td>
<td>110,895</td>
<td>106,632</td>
</tr>
</tbody>
</table>

Notes: This table shows the results from running equation 1 on the 1996-2009 period (columns 1-3) and the 2010-2015 period (columns 4-6). Specifications also include year dummies and dummies for different firm types: sole proprietors, partnerships and corporations. The donut hole, i.e. the excluded area, is defined to be the bunching region: 49700–54100 euro. *** p<0.01, ** p<0.05, * p<0.1.
A Depreciation Rules

The Finnish tax authorities’ definition of capital is any fixed assets which include all long-term tangibles that firms are using 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% to 20%, depending on the type of construction. For example, the annual depreciation rate for office buildings is 4%, 7% for factory buildings and 25% for immovable capital. The maximum rate of depreciation of machinery and equipment is 25%.

The life of assets can vary depending on the type of asset type that 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% of the costs of assets in a single tax year where the value for 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 CES and Magnitude of Bunching

Assume, as in previous sections that firms exhibit CES production functions $F(k, l) = (\alpha k^{\frac{1}{\sigma}} + (1 - \alpha)l^{\frac{1}{1-\sigma}})^{\frac{1}{1-\sigma}}$. Firms in the neighborhood of the payroll tax notch can either reduce their level of depreciation and bunch at the threshold in order to pay lower payroll taxes; or they remain above the threshold and pay higher payroll taxes. The relationship
between the magnitude of bunching and the capital-labor elasticity is pinned down by the “marginal buncher” i.e. the firm that is indifferent between bunching or paying the higher payroll tax rate. This firm’s profits are equal at the threshold and at the optimal point above it, since it is indifferent between bunching and not. Formally, denote by \( k_b \) the capital stock at the threshold and \( k_m \) the capital stock of this firm if it does not bunch.

**Above the threshold** If the marginal buncher locates above the threshold, then labor and capital are determined by the first order condition of the profit maximization problem and are related by the following equation:

\[
l = \left( \frac{r}{w} \frac{1 - \alpha}{\alpha} \right)^{\sigma} k_m
\]

Plugging this relationship in the profit function gives the profit function as a function of \( k_m \):

\[
\Pi(k_m) = k_m \left( p \left( \alpha + (1 - \alpha) \left( \frac{r}{w} \frac{1 - \alpha}{\alpha} \right)^{\sigma-1} \right)^{\sigma-1} - w \left( \frac{r}{w} \frac{1 - \alpha}{\alpha} \right)^{\sigma} - rl \right)
\]  
\[\text{(4)}\]

**At the threshold** At the threshold, the stock of capital is fixed and the firm is therefore maximizing profits with respect to labor holding capital fixed at \( k_b \):

\[
\max_l \Pi(k_b) = p \left( \alpha k_b^{\frac{\sigma-1}{\sigma}} + (1 - \alpha) l^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} - wl - r k_b
\]

The first order condition with respect to \( l \) is given by:

\[
l = k_b \left( \frac{1}{\alpha} \left( \frac{w(1 + \tau)}{p(1 - \alpha)} - (1 - \alpha) \right) \right)^{\frac{\sigma}{1-\sigma}}
\]

We plug this in the profit function to get:

\[
\Pi(k_b) = k_b \left( p \left( \alpha + \frac{1 - \alpha}{\alpha} \left( \frac{w(1 + \tau)}{p(1 - \alpha)} - (1 - \alpha) \right) \right)^{\frac{\sigma}{\sigma-1}} - w(1 + \tau) \left( \frac{1}{\alpha} \left( \frac{w(1 + \tau)}{p(1 - \alpha)} - (1 - \alpha) \right) \right)^{\frac{\sigma}{1-\sigma}} - r \right)
\]

Because the excess mass generated by bunching should be equal to the missing mass, the
distance between $k_m$ and $k_b$ is equal to the size of the excess mass i.e. $k_m = b + k_b$. The marginal buncher is indifferent between bunching or locating above the threshold which implies that $\Pi_m = \Pi_b$ and therefore:

$$b = k_b \left( \frac{A - B - C + D}{B - D - r} \right)$$

where $r$ is the cost of capital and

$$A = p \left( \alpha + \frac{1 - \alpha}{\alpha} \left( \frac{w(1 + \tau)}{p(1 - \alpha)} - (1 - \alpha) \right) \right)^{\frac{\sigma}{\sigma - 1}}$$

$$B = p \left( \alpha + (1 - \alpha) \left( \frac{r}{w} \frac{1 - \alpha}{\alpha} \right)^{\frac{\sigma}{\sigma - 1}} \right)^{\frac{\sigma}{\sigma - 1}}$$

$$C = w(1 + \tau) \left( \frac{1}{\alpha} \left( \frac{w(1 + \tau)}{p(1 - \alpha)} - (1 - \alpha) \right) \right)^{\frac{\sigma}{1 - \sigma}}$$

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

Equation \ref{eq:5} relates the size of the missing mass to the inverse of the capital-labor elasticity of substitution $\sigma$. Our goal is to use the difference in size of the missing mass across different industries and firm sizes to analyze the heterogeneity in their capital-labor elasticity of substitution. To do so, we analytically solve for the derivative of $b$ with respect to $\epsilon_{k,l} = \frac{1}{\sigma}$ and find that it increases for values of $\epsilon_{k,l}$ smaller than 1, which is the likely range over which firms are operating given our estimates that production functions are Leontief. Appendix Figure \ref{fig:8} shows this relationship.
Figure 8: Bunching and Elasticity

Notes: This Figure simulates equation 5 to show the relationship between the capital-labor elasticity of substitution and the magnitude of bunching.
<table>
<thead>
<tr>
<th>Year</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>I</th>
<th>II</th>
<th>I</th>
<th>II</th>
<th>Employees</th>
<th>Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>4.000</td>
<td>5.600</td>
<td>6.500</td>
<td>1.2</td>
<td>1.00</td>
<td>4.00</td>
<td>0.100</td>
<td>16.80</td>
<td>23.100</td>
<td>28.600</td>
</tr>
<tr>
<td>1997</td>
<td>4.000</td>
<td>5.600</td>
<td>6.500</td>
<td>1.4</td>
<td>1.00</td>
<td>4.00</td>
<td>0.090</td>
<td>16.70</td>
<td>23.190</td>
<td>28.690</td>
</tr>
<tr>
<td>1998</td>
<td>4.000</td>
<td>5.600</td>
<td>6.500</td>
<td>1.4</td>
<td>0.90</td>
<td>3.90</td>
<td>0.080</td>
<td>16.80</td>
<td>23.180</td>
<td>28.680</td>
</tr>
<tr>
<td>1999</td>
<td>4.000</td>
<td>5.600</td>
<td>6.500</td>
<td>1.3</td>
<td>0.90</td>
<td>3.85</td>
<td>0.080</td>
<td>16.80</td>
<td>23.080</td>
<td>28.530</td>
</tr>
<tr>
<td>2000</td>
<td>4.000</td>
<td>5.600</td>
<td>6.500</td>
<td>1.2</td>
<td>0.90</td>
<td>3.45</td>
<td>0.090</td>
<td>16.80</td>
<td>22.990</td>
<td>28.040</td>
</tr>
<tr>
<td>7/2000</td>
<td>3.600</td>
<td>5.600</td>
<td>6.500</td>
<td>1.2</td>
<td>0.90</td>
<td>3.45</td>
<td>0.090</td>
<td>16.80</td>
<td>22.590</td>
<td>28.040</td>
</tr>
<tr>
<td>2001</td>
<td>3.600</td>
<td>5.600</td>
<td>6.500</td>
<td>1.2</td>
<td>0.80</td>
<td>3.10</td>
<td>0.095</td>
<td>16.60</td>
<td>22.295</td>
<td>27.495</td>
</tr>
<tr>
<td>2002</td>
<td>3.600</td>
<td>5.600</td>
<td>6.500</td>
<td>1.1</td>
<td>0.70</td>
<td>2.70</td>
<td>0.095</td>
<td>16.70</td>
<td>22.185</td>
<td>27.085</td>
</tr>
<tr>
<td>3/2002</td>
<td>2.950</td>
<td>5.150</td>
<td>6.050</td>
<td>1.1</td>
<td>0.70</td>
<td>2.70</td>
<td>0.095</td>
<td>16.70</td>
<td>21.535</td>
<td>26.635</td>
</tr>
<tr>
<td>2003</td>
<td>2.964</td>
<td>5.164</td>
<td>6.064</td>
<td>1.1</td>
<td>0.60</td>
<td>2.45</td>
<td>0.081</td>
<td>16.80</td>
<td>21.545</td>
<td>26.495</td>
</tr>
<tr>
<td>2004</td>
<td>2.964</td>
<td>5.164</td>
<td>6.064</td>
<td>1.1</td>
<td>0.60</td>
<td>2.50</td>
<td>0.080</td>
<td>16.80</td>
<td>21.544</td>
<td>26.544</td>
</tr>
<tr>
<td>2005</td>
<td>2.966</td>
<td>5.166</td>
<td>6.066</td>
<td>1.2</td>
<td>0.70</td>
<td>2.80</td>
<td>0.080</td>
<td>16.80</td>
<td>21.746</td>
<td>26.946</td>
</tr>
<tr>
<td>2006</td>
<td>2.958</td>
<td>5.158</td>
<td>6.058</td>
<td>1.1</td>
<td>0.75</td>
<td>2.95</td>
<td>0.080</td>
<td>16.70</td>
<td>21.588</td>
<td>26.888</td>
</tr>
<tr>
<td>2007</td>
<td>2.951</td>
<td>5.151</td>
<td>6.051</td>
<td>1.1</td>
<td>0.75</td>
<td>2.95</td>
<td>0.080</td>
<td>16.64</td>
<td>21.521</td>
<td>26.821</td>
</tr>
<tr>
<td>2008</td>
<td>2.771</td>
<td>4.971</td>
<td>5.871</td>
<td>1.0</td>
<td>0.70</td>
<td>2.90</td>
<td>0.080</td>
<td>16.80</td>
<td>21.351</td>
<td>26.651</td>
</tr>
<tr>
<td>2009</td>
<td>2.801</td>
<td>5.001</td>
<td>5.901</td>
<td>1.0</td>
<td>0.65</td>
<td>2.70</td>
<td>0.070</td>
<td>16.80</td>
<td>21.321</td>
<td>26.471</td>
</tr>
<tr>
<td>4/2009</td>
<td>2.000</td>
<td>4.201</td>
<td>5.101</td>
<td>1.0</td>
<td>0.65</td>
<td>2.70</td>
<td>0.070</td>
<td>16.80</td>
<td>20.520</td>
<td>25.601</td>
</tr>
<tr>
<td>2010</td>
<td>2.220</td>
<td>2.220</td>
<td>2.220</td>
<td>0.8</td>
<td>0.75</td>
<td>2.95</td>
<td>0.070</td>
<td>16.90</td>
<td>20.740</td>
<td>22.940</td>
</tr>
<tr>
<td>2011</td>
<td>2.210</td>
<td>2.210</td>
<td>2.210</td>
<td>1.0</td>
<td>0.80</td>
<td>3.20</td>
<td>0.070</td>
<td>17.10</td>
<td>21.180</td>
<td>23.580</td>
</tr>
<tr>
<td>2012</td>
<td>2.210</td>
<td>2.210</td>
<td>2.210</td>
<td>1.0</td>
<td>0.80</td>
<td>3.20</td>
<td>0.070</td>
<td>17.35</td>
<td>21.430</td>
<td>23.830</td>
</tr>
<tr>
<td>2013</td>
<td>2.040</td>
<td>2.040</td>
<td>2.040</td>
<td>0.9</td>
<td>0.80</td>
<td>3.15</td>
<td>0.070</td>
<td>17.35</td>
<td>21.160</td>
<td>23.510</td>
</tr>
<tr>
<td>2014</td>
<td>2.140</td>
<td>2.140</td>
<td>2.140</td>
<td>0.9</td>
<td>0.75</td>
<td>2.95</td>
<td>0.070</td>
<td>17.75</td>
<td>21.610</td>
<td>23.810</td>
</tr>
<tr>
<td>2015</td>
<td>2.080</td>
<td>2.080</td>
<td>2.080</td>
<td>0.9</td>
<td>0.80</td>
<td>3.15</td>
<td>0.070</td>
<td>18.00</td>
<td>21.850</td>
<td>24.890</td>
</tr>
<tr>
<td>2016</td>
<td>2.120</td>
<td>2.120</td>
<td>2.120</td>
<td>0.8</td>
<td>1.0</td>
<td>3.90</td>
<td>0.070</td>
<td>18.00</td>
<td>21.990</td>
<td>24.890</td>
</tr>
<tr>
<td>2017</td>
<td>1.080</td>
<td>1.080</td>
<td>1.080</td>
<td>0.8</td>
<td>0.8</td>
<td>3.30</td>
<td>0.070</td>
<td>17.95</td>
<td>20.700</td>
<td>23.200</td>
</tr>
</tbody>
</table>

Table 3: Social insurance percentages by firm categories, different insurance types and years

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