Can compulsory staff registers reduce tax evasion? Results from a Swedish reform

Sven-Olov Daunfeldt[*], Anton Gidehag[†] and Niklas Rudholm[‡]

Abstract

It has long been realized that firms can evade taxes by underreporting their wage payments. In 2007, the Swedish government tried to prevent such behavior by implementing a reform that required restaurants and hairdressers to have staff registers with detailed information on when their employees were working. The Swedish Tax Authority was given a mandate to carry out unannounced control visits and impose fines on those firms that had not properly filled out their staff register. We identify control industries that are similar to the treated industries using propensity score matching, and then investigate the effect of staff registers on wages per employee using a firm-level difference-in-difference regression model. We find no significant effects of staff registers on wage payments up to three years after the staff register reform was implemented. Compulsory staff registers thus seem as an inefficient policy instrument to reduce unreported wages, especially considering that the reform is associated with administrative costs and increased regulatory burden for the treated firms.

Keywords: tax evasion, firm regulation, quasi-experimental method, unreported wages, propensity score matching

JEL classifications: H26, H32, K34, L51

[*] Corresponding author. E-mail: sven-olov.daunfeldt@huiresearch.se. Telephone: +46702957284. Institute of Retail Economics, Stockholm, Sweden and Dalarna University, Falun, Sweden.
[†] Institute of Retail Economics, Stockholm, Sweden and Örebro University, Örebro, Sweden.
[‡] Institute of Retail Economics, Stockholm, Sweden.

1. Introduction

Tax evasion is a major problem in many countries, with some evidence suggesting that noncompliance among firms is as problematic as individual noncompliance (Fisher and Goddeeris 1988). However, the literature on tax evasion has mainly been concerned with individuals' decision to evade taxes and offers little guidance on appropriate policy actions for reducing tax evasion among firms (Joulfaian 2000; Crocker and Slemrod 2005).[1]

One of the most common ways for employers to evade taxes is to deliberately understate their wage payments (Yaniv 1988), and policymakers can then respond by enhancing the tax authorities' ability to detect and punish such behavior. This happened in Sweden on January 1, 2007, when the policymakers implemented a law that required restaurants and hairdressers to have a staff register. Employers were required to provide detailed information on when their employees came and left the workplace, and the law also allowed the Swedish Tax Authority to make unannounced control visits to ensure that the staff registers were filled out correctly. If not, the firm was required to pay a fixed amount of 10,000 Swedish krona (SEK), approximately $1,126,[2] and an additional 2 000 SEK ($225) for each individual for whom there was inadequate information. In 2007, more than 31 000 control visits were carried out and 3 515 firms had to pay fines because they had not filled in their staff registers properly (Swedish Tax Authority 2015).

Although the reform made it easier for the tax authorities to detect tax evasion, it also increased firms' regulatory burden. The total yearly administrative costs for the firms that needed to introduce staff registers in 2007 have been estimated to be 365 million SEK (46 million USD) (NUTEK 2008). The reform also led to increased costs for the tax authorities since they needed

[1] Note that it exists some theoretical studies on business tax evasion (e.g., Marrelli 1984; Wang and Conant 1988), while empirical studies have largely ignored tax evasion among firms.

[2] Based on the exchange rate January 11, 2019.

additional resources to carry out the controls and to administer the staff register system.[3] Any increase in tax revenue due to the staff register requirement must at least exceed these direct costs of the reform to be beneficial for the taxpayers. Although outside the scope of the present paper, it should be noted that the reform might also change the behavior of firms in ways that are non-beneficial for society. The introduction of staff registers increases, for example, firms' regulatory burden, which tend to have a negative influence on growth ambitions, productivity, and the number of start-ups (Klapper et al. 2006; Schivardi and Viviano 2011; Dreher and Gassebner 2013).

Our aim is to investigate the impact of the 2007 Swedish introduction of staff registers on firms' reported wages, and whether this reform can be justified on economic grounds. The Swedish Tax Authority (2009) has previously found that the introduction of staff registers increased reported total wages of the treated firms compared to years before the reform was implemented. However, their report does not consider that the total wage sums in the restaurant and hairdresser industries could have increased over time even in absence of the staff register requirement. Another problem is that firms' payroll taxes also was reduced by 11 percentage points for all employees that were between 19-25 years old in 2007. Restaurants and hairdressers have substantially larger shares of young employees than most other Swedish industries (RAMS register, Statistics Sweden 2017), and previous studies show that firms increased their number of young employees due to the payroll tax cut (Egebark and Kaunitz, 2014). Total wage sums thus increased at the time of the reform in the targeted industries independently of the introduction of staff registers, which implies that the Swedish Tax Authority (2009) evaluation captures the joint effect of two reforms introduced simultaneously.

We choose to investigate the effects of the introduction of staff registers on the wages per employee instead of total wage sums because the former measure is less likely to be influenced by

[3] We are not aware of any estimations of these costs.

the payroll tax reform in 2007.[4] Furthermore, to address that the wages per employee might have changed more in the targeted industries even in absence of staff registers, we use an empirical method that identifies control industries that have similar shares of young employees and trends in wage reporting as the treated industries during the pre-reform years. Our reasoning is as follows. If the treated and control group industries have had similar shares of young employees and trends in wage reporting before the reform, it is likely that the payroll tax reform affected these industries in a similar manor, and that the trends in wages would have been similar in the post-reform years if the staff register reform had not been implemented.

More specifically, we use propensity score matching (Rosenbaum and Rubin 1983) to identify control industries with similar pretreatment characteristics regarding the share of young employees and trends in the dependent variable as the treated industries. We then estimate the effect of staff registers on the reported wages per employee by comparing the outcomes for firms in the treatment and control groups using firm-level difference-in-difference estimations. Our analysis is based on matched employer-employee data from Statistics Sweden covering all Swedish residents that are at least 16 years old.

We find that wage payments are not particularly sensitive to the introduction of staff registers, and we cannot reject the null hypothesis that the staff register requirement had no effect on wages per employee up to three years after its implementation. After four years, we find that wages per employee in the restaurant industry increased by an average of 2.12 percent more than in the control group industries. However, the long-term estimates are less likely to be due to the reform since there is an increased likelihood of events affecting control and treatment industries differently as time goes by. We conclude that the introduction of staff registers has been an inefficient method

[4] The wages per employee might also be affected by the payroll taxes if it affects the composition of the workforce. As a robustness check, we therefore also investigate the effect of staff registers on the wages of employees that worked at the firm both before and after the introduction of staff registers.

of reducing unreported wages, especially considering the costs of performing control visits, the administrative costs imposed on the firms, and the possible negative effects on entrepreneurial behavior.

The article is organized as follows. Section 2 describes the data and our empirical approach, including our identification of control industries. The regression model and our estimation results are presented in section 3. Finally, our results are summarized and discussed in section 4.

2. Data and Matching Model

2.1. Data

Our empirical analysis is based on the individual-level database LISA (Longitudinal Integration Database for Health Insurance and Labour Market Studies), which is provided by Statistics Sweden. LISA is a database that is built upon a number of different registers and it includes information on all Swedish residents that are at least 16 years old.

The RAMS register (Labour Statistics Based on Administrative Sources) provides information on individuals' employment status and potential employers in November each year. The data also includes an identification number for the employer, which means that we can match all employed individuals with their respective employer. We are thus able to construct a panel of Swedish firms and their employees from 2003 to 2010. It includes information on the number of employees and total gross wages, making it possible for us to calculate wages per employee at firm *i* in year *t*.

All firms are assigned industry codes that reveal their industrial affiliation. The industry codes are derived from the SNI2002 (Swedish Standard Industrial Classification) classification system.[5] SNI2002 consists of 17 industry groups at the most aggregated level, and 776 industry

[5] For more information about SNI2002, see http://www.scb.se/en_/Documentation/Classifications-and-standards/Swedish-Standard-Industrial-Classification-SNI/

groups at the most detailed five-digit level. To be able to identify representative control industries using the propensity score matching method, we construct a panel at the five-digit industry level containing measures of, for instance, gross wages.

The total gross wages for each firm and industry are derived by exploiting workers' annual gross wages from their primary employers.[6] More specifically, for each firm observed during the month of November, we aggregate the gross wages for workers having that firm as their primary income source (their primary workplace). Hence, we will underestimate a firm's total gross wages if there are workers registered at the firm in November that have another primary employer.[7]

We restrict our sample to firms having at least two employees. The reason is that most one-employee firms within the targeted industries were exempted from the staff register requirement.[8]

2.2 Identification strategy

The fundamental identification problem that we need to address is that we cannot observe those firms that were required to introduce staff registers in the counterfactual state of not being subject to the reform. The targeted industries were not randomly chosen, but rather because the policymakers believed that firms within these industries were especially prone to evade taxes by underreporting wages. In order to identify the effect of staff registers on reported wages, we must

[6] More specifically, we use the LISA variable 'KU1Ink'. Note that the income on a yearly basis must exceed 1000 SEK to be reported. The variable is measured in 100 SEK. See Statistics Sweden (2016).

[7] Approximately 80 percent of the workers were employed at their primary employer during the month of November. The labor mobility is slightly higher in the restaurant sector, suggesting slightly more underestimated wages. However, the outcome variable (wages per employee) is based solely on employees working at their primary workplace in November.

[8] More specifically, independent contractors, closely held companies and closely held partnerships in which only the chief executive or his/her family is active are exempted. Moreover, firms whose main operation is within a non-targeted industry are also exempted. The rule for being exempted is that at least 75 percent of the firm turnover is associated with a non-targeted industry.

account for the counterfactual outcome of how wage reporting in these industries would have developed if staff registers had never been implemented.

The true counterfactual outcome is impossible to observe for obvious reasons, but different statistical methods have been developed to find measures of counterfactual outcomes. We use a two-step method to estimate the effect of staff registers on reported wages. First, as treatment is directed toward industries rather than individual firms, we use propensity score matching (Rosenbaum and Rubin 1983) to identify control industries that had similar trends in wages per employee and similar shares of young employees as the treated industries in the years leading up to the reform. The first matching criterion is important since similar (parallel) trends is the basis for our second step difference-in difference analysis. Having similar trends in wage reporting across treated and control industries in the years leading up to the reform makes it probable that treated industries would have had similar trends in wage reporting as the control industries in the post-reform years if staff registers had not been implemented. The latter matching criterion is of importance to isolate the effect of staff registers from the potential impact of the youth payroll tax cut carried out in 2007, and industries with similar shares of young employees should have been affected similarly by the payroll tax reform when it comes to its effect on reported wages.

In the next step, we estimate the treatment effect by comparing the development of reported wages per employee among firms in the treated industries (i.e., restaurants and hairdressers) and firms in the matched control industries. More specifically, we estimate a difference-in-difference model to compare pre- and posttreatment changes between treated and control group firms (Card and Krueger 1994; Abadie 2005; Angrist and Pischke 2008).

2.3 Identification of control industries

To obtain propensity scores for the five-digit industries under study, i.e., their estimated probability of getting treatment, we estimate the following probit model:

$$Treated_{j,t=07} = \alpha + \beta_1 Wage/emp_{j,t=06} + \beta_2 \Delta Wage/emp_{j,t=05} + \beta_3 \Delta Wage/emp_{j,t=04} + \beta_4 Size_{j,t=06} + \beta_5 Firm\ size_{j,t=06} + \beta_6 Share\ young_{j,t=06} + \beta_7 Wage/emp^2_{j,t=06} + \beta_8 \Delta Wage/emp^2_{j,t=05} + \beta_9 \Delta Wage/emp^2_{j,t=04} + \beta_{10} Size^2_{j,t=06} + \beta_{11} Firm\ size^2_{j,t=06} + \beta_{12} Share\ young^2_{j,t=06} + \varepsilon_{jt}, \quad (1)$$

where the dependent variable $Treated_{j,t=07}$ is a variable that is equal to one for all SNI2002 five-digit restaurant and hairdresser industries and zero for all other industries during the treatment year.[9] We control for the past wage development by including wages per employee in the five-digit industry j in 2006, $Wage/emp_{j,t=06}$; its annual growth over the years 2003-2005, represented by $\Delta Wage/emp_{j,t=05}$ and $\Delta Wage/emp_{j,t=04}$; total number of employees within each industry in 2006, $Size_{j,t=06}$; the average firm size within each industry in 2006, $Firm\ size_{j,t=06}$; and the share of employees in each industry that is younger than 26 years in 2006, $Share\ young_{j,t=06}$. The latter variable accounts for the potentially heterogeneous effect of the 2007 youth payroll tax cut. As suggested by Angrist and Pischke (2008), we also include all variables in their squared forms to control for non-linear relationships. Thus, we seek to identify control industries with similar wage development, labor force composition and size as the restaurant and hairdresser industries in the years before the staff register requirement was implemented.[10]

From the estimation of equation (1), we obtain propensity scores for all five-digit industries. Using nearest neighbor matching, we then assign the five industries with most similar propensity scores to each five-digit restaurant and hairdresser industry. The treated restaurant and hairdresser

[9] The regression results are presented in Appendix 1 (Table A1).
[10] It should also be noted that controlling for past wage development implies partly controlling for the economic development within each industry. Hence, we at least partly account for potential differences due to industry-specific economic shocks.

industries and their corresponding control industries are presented in Table A2 in the Appendix. In total, we identify 25 unique control industries at the five-digit SNI2002 code level that have similar propensity scores as the restaurant and hairdresser industries.

Our matching model assigns 8 891 and 1 222 control firms to 16 340 and 2 865 firms within the restaurant and hairdresser industries, respectively. In order to ensure that our matching variables have similar characteristics within the treatment and control groups, we also perform a balancing test. We have managed to identify industries that constitute a valid control group if the underlying variables that affect treatment assignment have similar characteristics, implying that treatment should be *as if* randomly assigned between the treated and matched control industries. The results of our balancing test are presented in Table 1, showing that the means of the matching variables are not significantly different between the treated and control industries after matching. The average one-year lagged wage sum and the share of youth employees are, for example, very similar. These findings indicate that we have found valid control groups for the treated industries.

[Table 1 about here]

2.4 Descriptive statistics

Table 2 presents descriptive statistics for firms operating within the targeted industries and their control industries. We also include all other firms for comparison.

[Table 2 about here]

From Table 2, it is apparent that firms within the control industries are noticeably more similar to the treated firms than other (unmatched) firms. Considering the total wages per firm, they are on average considerably lower among restaurant and hairdresser firms. The average and median wages per employee are also lower among firms within these industries. These wage differences are also reflected in the average firm size. The average restaurant has nearly seven employees, whereas

the average hairdresser has almost four employees. The corresponding numbers for control firms and other firms are 18 and 24 employees, respectively.

3. Regression model

Difference-in-difference analysis rests on the assumption of parallel trends in the outcome variable among treated and control units in the posttreatment period in the absence of treatment, implying that they would have had identical outcomes if the treatment had never been implemented.[11] This assumption is commonly investigated by examining whether the pre-treatment trends are parallel (Angrist and Pischke 2008; Ryan et al. 2015). The average logged wages per employee among the restaurant and hairdresser firms and their respective control firms in the pretreatment years are therefore presented in Figure 1.[12]

[Figure 1 about here]

Figure 1 shows that the average wages are lower for firms in the restaurant industry compared to their control group of firms, while the opposite is true for hairdressers. However, it is not important that the levels are similar in difference-in-difference estimations, but rather that the trends are parallel. As can be seen from Figure 1, we observe similar trends for all pretreatment years for restaurant firms and their controls. The pre-treatment trends in wages are also similar for hairdressers and control firms during 2004-2006, but they differ somewhat during 2003-2004. Note, however, that these differences are small in practice since wages are expressed in logged values and presented on a fine scale.

[11] Note that this is also related to our choice of dependent variable. If using the total wages of each firm in the analysis, the identifying assumption is that the payroll tax reform that was implemented at the same time as the staff register reform identically affected employment and wages in the treatment and control group industries. By using wages per employee instead of total wages, we allow for the possibility that the payroll tax reform had different impacts on employment in our treatment and control group industries.

[12] Throughout the paper, we do exclude firms with extreme vales in the outcome variable. A firm is defined as an outlier if the annual growth in our outcome variable (wages per employee) deviates by more than three standard deviations from the average annual growth in wages per employee. We exclude 4 610 firms out of 629 662 firms during 2006-2010 (or 16 748 out of 2 022 590 observations).

In order to investigate how the introduction of staff registers affected wage reporting in the treated firms, we estimate the following firm-level difference-in-difference model:

$$\ln Y_{it} = \alpha + \gamma TI_i + \lambda TP_t + \sigma(TI_i * TP_t) + \eta_i + \varepsilon_{it} \qquad (2)$$

where the dependent variable $\ln Y_{it}$ is the natural logarithm of the wages per employee at firm i in year t. There are two reasons for expressing the outcome variable in the log-form. First, the variable becomes approximately normally distributed, which is good for statistical inference. Second, it yields a semi-elastic model in which the estimated treatment effects can be interpreted as percentage changes.[13]

TI_i is a treatment indicator that is equal to one for firms within the treatment industries, and equal to zero for firms within the matched control industries. The treatment indicator controls for potential level differences in wages per employee between the firms in the treatment and control group. TP_t is an indicator that takes the value of one during the treatment period (2007-2010) and zero in the year prior to treatment (2006).[14] We choose to use only 2006 as our pre-intervention year since the trends in Figure 1 show a strong similarity in the last year leading up to the reform but somewhat larger differences in previous years, especially considering hairdresser firms.[15] By including TP_t, we control for time-variant effects that are common for both treated and control firms. For instance, it captures the general economic factors affecting the development of wages per employee within all industries.

Our main variable of interest is the interaction term between TI_i and TP_t, which is equal to 1 for the treated firms during the treatment period. Its parameter σ can be expressed as follows:

[13] The exact effect in percentage terms of a parameter estimate σ can be calculated using the formula 100 × [exp(σ) – 1]. However, since the parameter estimates in our setting are small, the differences are negligible.

[14] Notice that since all data is collected in November each year, year 2007 is technically a post-treatment year.

[15] Another reason for only including one pre-treatment year is that the estimated reform effect becomes less precise the further away one moves from the reform introduction. Put differently, the estimated reform effect is most accurate close to the year of its introduction (Mian and Sufi 2012).

$$\sigma = E[ln\, Y_{it}|Tr_i = 1, t = Post] - E[ln\, Y_{it}|Tr_i = 1, t = Pre] -$$

$$(E[ln\, Y_{it}|Tr_i = 0, t = Post] - E[ln\, Y_{it}|Tr_i = 0, t = Pre])$$

Thus, parameter σ represents the differences in the conditional means within the treatment and control groups, before and after treatment. Consequently, our estimated parameter $\hat{\sigma}$ compares how the average logged wages per employee have changed within the treatment and control group firms at the time of the reform.

To ensure that our results are not driven by time-invariant firm-specific heterogeneity among firms in the intervention and control groups, we also include firm-specific random effects η_i.[16] Lastly, ε_{it} is an idiosyncratic error term.

4. Results

4.1. Main results

The results when estimating equation (2) are presented in Figures 2-4, with the point estimates and their 95 % confidence intervals highlighted. Estimates of the treatment effect is statistically significantly different from zero with 95 percent certainty if the confidence interval does not cross the x-axis at zero. We choose to only include firms with at least two employees to be certain that only the firms that are required to fill out staff registers are included in the analysis. We estimate the effect of staff registers up to four years after the reform was introduced, i.e. until 2010. The full regression results are presented in Tables A3-A5 in the Appendix.[17]

We start by presenting the estimated effect of the staff registers on wages per employee for restaurants and hairdressers jointly in Figure 2, followed by the separate estimated effects for each

[16] Due to higher efficiency, the firm specific (time invariant) heterogeneity is accounted for as random effects rather than as fixed effects. However, using fixed effects yields very similar point estimates. These results are available upon request.

[17] The estimates in the figures correspond to the fourth column of each table.

industry. The results in Figure 2 show that the estimated effects of the staff register requirement on wages per employee are -0.76 and -0.49 percent during the first and second post-reform years, respectively. Neither estimate is statistically significantly different from zero, which means that we cannot reject the null hypothesis that the introduction of staff registers had no effect on reported wages per employee in the short-run.

[Figure 2 about here]

We only obtain a positive and significant estimate four years after the reform, suggesting that the introduction of staff registers increased the reported wages per employee by an average of 1.51 percent. However, the long-term estimates are less likely to be due to the reform since there is an increased likelihood of events affecting treatment and control industries differently. The financial crisis that reached Sweden in 2009 could, for example, impact the treated and control industries in different ways, biasing the estimation of the treatment effect.

In Figure 3, we present the results when estimating the impact of staff registers on wages per employee within the restaurant industry separately. Again, we cannot observe any significant effects of staff registers on reported wages during the first three post-reform years. Meanwhile, a positive and significant effect is found after four years.

[Figure 3 about here]

The results for hairdresser firms are presented in Figure 4. In this case, neither the short-run nor the long-run estimates are significantly different from zero. Thus, staff registers do not seem to have any impact on reported wages within the hairdresser industry.

[Figure 4 about here]

To summarize, we find no evidence that the staff register requirement had a positive effect on firms' short-run wage reporting. In the long run, our results indicate that the reform increased wage reporting among restaurants. However, these long-run estimates should be interpreted with caution since other factors, such as the financial crisis, might have affected the estimates through a heterogeneous impact on the wage development in the reform and control group industries.

4.2. Robustness checks

As previously discussed, the payroll tax cut for young employees was carried out parallel to the staff register reform, and could thus bias our estimates. We have tried to control for this by including the share of young employees in the matching model and investigate the effect of staff registers on wages per employee instead of total wages. However, although the wages per employee is less likely to be a biased measure, the payroll tax reform could still generate biased estimates if:

i) The payroll tax cut leads to wage increases for already employed workers. This can be the case because a reduction of labor costs typically leads to a spillover to employees through wage increases (see for instance Cruces et al. 2010) If so, our estimates of the effect of staff registers are *overestimated.*

ii) The payroll tax cut increases the employment of youths whom generally have below average wages. Such an effect would reduce the average wage per employee, causing our estimates to be *underestimated.*

We do not know which of these two conflicting potential sources of bias that dominates. As a robustness check, we therefore choose to only analyze the wage development among individuals who were employed at the same firm both before and after the reform. This means that we eliminate any potential bias from ii), implying that the estimated effects of the reform on wages per employee only can be upward biased due to i).

For this robustness analysis, we redo our matching procedure using the same set of variables as above to identify control industries with similar wage developments for individuals staying at the

same workplace for at least two consecutive years in the pre-reform period.[18] Next, we perform difference-in-difference analysis at the firm level by comparing wage changes for individuals staying at either the same treated or control industry firm before and after the implementation of staff registers. These results can be found in the appendix (Tables A6 and A7).

We find a positive and significant reform effect for the restaurants, suggesting that the staff register requirement increased wages for incumbent workers by an average of 2.17 and 6.55 percent one and four years after the reform, respectively. Hence, the staff register requirement appears to have led to increased wage reporting among individuals employed by the same restaurant firm during the entire period of study. However, considering that the payroll tax reform likely resulted in wage spillovers, we also consider these estimates to be overestimated. For hairdresser firms, we find no evidence of a short-term or long-term link between staff registers and wage reporting.

5. Summary and discussion

Wage underreporting among firms can have significant negative impacts on government revenues. One possibility to prevent such behavior is to implement reforms that makes it easier for the tax authorities to detect and punish firms that underreport their wage payments. Such a reform was implemented in Sweden in 2007, requiring restaurants and hairdressers with at least two employees to introduce staff registers. The reform also allowed the Swedish Tax Authority to make unannounced control visits, and firms were required to pay substantial fines in case of misreporting.

We have investigated the efficiency of this reform by first creating a control group of firms that were active in industries with similar wage development and shares of young employees as firms in the treated industries during the pre-treatment period. We then compare how wages per employee

[18] The balancing test and the matched control industries can be found in Tables A8 and A9.

evolved pre-reform and post-reform for firms in the treatment and control groups by estimating a firm-level difference-in-difference regression model.

We found no statistically significant effects of staff registers on wages per employee during the first three post-reform years in our main model specification. Our results were robust to numerous alternative specifications, implying that staff registers are an inefficient way of increasing wage reporting among firms. The largest significant positive one-year estimate of the reform effect implied that wages per employee increased by 2.17 percent in the restaurant industry (Table A6), and we know that this estimate is likely to overestimate the effect of the reform.

However, even in this case, we argue that the estimated effect is only statistically rather than economically significant. According to the Swedish Tax Authority (2009, p. 248-251), there were 14 958 restaurants with total wages of 13.1 billion SEK (1.47 billion USD) in 2007 (corresponding to 54 189 full-time employees). The average wages per restaurant firm was thus 875,786 SEK ($98,580), which was shared by 3.62 full-time employees. If we assume that each firm within the restaurant industry increased their wages per employee by 2.17 percent, total wages would increase by 284 million SEK (32 million USD) (0.0217*875 786*14 958). Assuming that the average wage tax rate is 50 %, this would result in an increase in tax revenues by 142 million SEK (16 million USD).[19] This can be compared with the annual administrative costs that have been estimated to be 365 million SEK (46 million USD) (NUTEK 2008). The increase in tax revenues is thus smaller than even the firms' administrative costs associated with the reform.

What can then explain the lack of positive significant effects of the reform on tax revenues? One possible explanation is that only a small number of the targeted firms evade taxes from the beginning. This could explain why our estimates representing the change in conditional means are

[19] The Swedish Government (2017) estimates the average tax rate on labor income (including social security fees) to be 48.3 percent

small and mostly non-significant. Another explanation could be that firms find ways to circumvent the staff register requirement.

The Swedish Tax Authority (2009) has previously concluded that the staff register reform has been successful in reducing unreported employment. Despite methodological shortcomings and extensive critiques of the study from a number of organizations and decision-making bodies (Board of Swedish Industry and Commerce for Better Regulation 2010; Confederation of Swedish Enterprise 2010; The Swedish Better Regulation Council 2010), the Swedish policymakers have introduced staff registers within the laundry and construction industries in 2013 and 2016, respectively. The Swedish government has recently also implemented an extension of the staff register requirement to include industries such as vehicle repair, beauty care and wholesale of food to be implemented January 1, 2019 (Swedish Government 2017). Our findings indicate that it is highly doubtful if such an extension of the staff register requirement will reduce unreported wages.

We conclude that the staff register system has been an inefficient way of increasing reported wages. The reform has also increased the regulatory burden for firms, and is therefore likely to induce indirect costs that are difficult to measure. We believe that more research on tax evasion among firms in general is needed, and particularly on how firms respond to different institutional reforms that are supposed to reduce tax evasion.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, 72(1), 1-19.

Angrist J.D., & Pischke, J.S. (2008). *Mostly harmless econometrics: an empiricist's companion*. Princeton, NJ: Princeton University Press

Board of Swedish Industry and Commerce for Better Regulation. (2010). Angående remiss av departementspromemorian Närvaroliggare och kontrollbesök Ds 2009:43. Stockholm

Card, D., & Krueger, A. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4), 772-793.

Confederation of Swedish Enterprise. (2010). Remissyttrande – Departementspromemorian Närvaroliggare och kontrollbesök Ds 2009:43. Stockholm

Crocker, K.J., & Slemrod, J. (2005). Corporate tax evasion with agency costs. *Journal of Public Economics*, 89(9), 1593-1610.

Cruces, G., Galiani, S., & Kidyba, S. (2010). Payroll taxes, wages and employment: Identification through policy changes. *Labour Economics*, 17(4), 743-749.

Dreher, A., & Gassebner, M. (2013). Greasing the wheels? The impact of regulations and corruption on firm entry. *Public Choice*, 155 (3-4), 413-432.

Egebark, J., & Kaunitz, N. (2014). Do payroll tax cuts raise youth employment?. *IFN Working Paper No. 1001*, Research Institute of Industrial. Economics, Stockholm.

Fisher, R.C., & Goddeeris, J.H. (1988). Participation in state tax amnesties: the case of business taxes. In Proceedings of the 81st annual conference on taxation held under the auspices of the National Tax Association-Tax Institute of America (pp. 139-145). National Tax Association-Tax Institute of America, Columbus, OH.

Joulfaian, D. (2000). Corporate income tax evasion and managerial preferences. *The Review of Economics and Statistics*, 82(4), 698-701.

Klapper, L., Laeven, L., & Rajan, R. (2006). Entry regulation as a barrier to entrepreneurship. *Journal of Financial Economics*, 82(3), 591-629.

Marrelli, M. (1984). On indirect tax evasion. *Journal of Public Economics*, 25(1-2), 181-196.

Mian, A., & Sufi, A. (2012). The effects of fiscal stimulus: Evidence from the 2009 cash for clunkers program. *The Quarterly Journal of Economics*, 127(3), 1107-142.

NUTEK. (2008). Näringslivets administrativa kostnader för skatteområdet – uppdatering 2007 R 2008:42. Stockholm.

Rosenbaum, P.R., & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.

Ryan, A.M., Burgess, J.F., & Dimick, J.B. (2015). Why we should not be indifferent to specification choices for difference-in-differences. *Health Services Research*, 50(4), 1211-1235.

Schivardi, F., & Viviano, E. (2011). Entry barriers in retail trade. *The Economic Journal*, 121(551), 145-170.

Statistics Sweden. (2016). Longitudinell integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier (LISA) 1990–2013. Documentation. http://www.scb.se/Statistik/AM/AM9901/_dokument/AM9901_1990I13_BR_AM76BR1601.pdf. Accessed January 29, 2019

Statistics Sweden. (2017). Registerbaserad arbetsmarkadsstatistik/Labour statistics based on administrative sources register. Overview. https://www.scb.se/hitta-statistik/statistik-efter-amne/arbetsmarknad/sysselsattning-forvarvsarbete-och-arbetstider/registerbaserad-arbetsmarknadsstatistik-rams/produktrelaterat/Fordjupad-information/beskrivning-av-registerbaserad-arbetsmarknadsstatistik-rams/. Accessed January 29, 2019

Swedish Government. (2017). Lagrådsremiss – Personalliggare i fler verksamheter, Stockholm.

Swedish Tax Authority. (2009). Skapar personalliggare fler vita jobb? En utvärdering av personalliggarnas effekter på skatteundandragande i restaurang- och frisörbranscherna. Bilaga 4. In: Närvaroliggare och kontrollbesök – en utvärdering och förslag till utvidgning, Ds 2009:43, Ministry of Finance, 239–348.

Swedish Tax Authority. (2015). Skatter i Sverige – Skattestatistisk Årsbok 2015. Solna, Stockholm.

The Swedish Better Regulation Council. (2010). Yttrande över promemorian Närvaroliggare och kontrollbesök (Ds 2009:43), Stockholm.

Wang, L.F., & Conant, J.L. (1988). Corporate tax evasion and output decisions of the uncertain monopolist. *National Tax Journal*, 41(4), 579-581.

Yaniv, G. (1988). Withholding and non-withheld tax evasion. *Journal of Public Economics*, 35(2), 183-204.

Tables and Figures (To be included in paper)

Table 1 Balancing test for treated and control industries

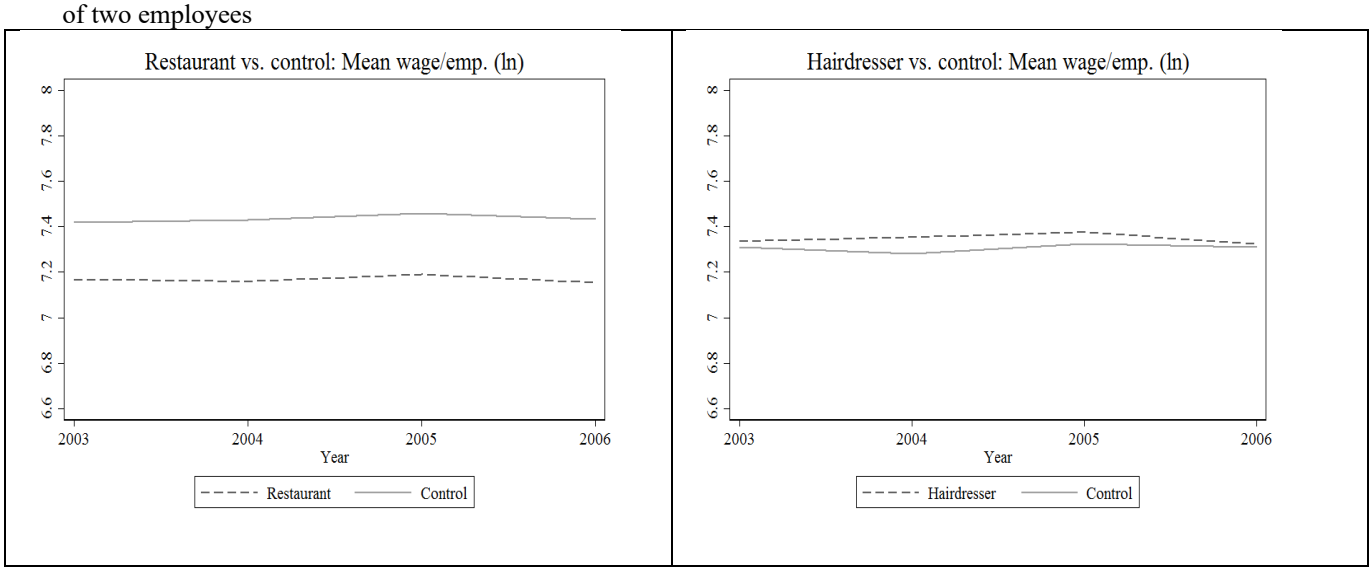
Variables	B/A = Before/After matching	Mean		T-test		
		Treated	Control	Bias (%)	t	p-value
Wage/emp., t=2006	B	1847.9	2713.5	-144	2.89	0.004
	A	1847.9	1791.6	9.4	0.39	0.702
(Wage/emp.)^2, t=2006	B	3.50e+06	8.00E+06	-130.4	-2.51	0.012
	A	3.50e+06	3.20E06	7.3	0.48	0.638
ΔWage/emp., t=2005	B	52.796	86.476	-16.9	-0.32	0.752
	A	52.796	52.404	0.2	0.04	0.970
(ΔWage/emp.)^2, t=2005	B	3104.3	86873	-13.4	-0.25	0.802
	A	3104.3	3040.3	0	0.06	0.953
ΔWage/emp., t=2004	B	-17.497	37.56	-29.6	-0.68	0.498
	A	-17.497	12.797	-16.3	-0.49	0.634
(ΔWage/emp.)^2, t=2004	B	20200	47247	-9.5	-0.18	0.858
	A	20200	3349.9	5.9	0.97	0.350
Size, t=2006	B	11141	4884.5	28.5	0.71	0.481
	A	11141	8511.9	12	0.3	0.771
(Size)^2, t=2006	B	4.80E+08	5.70E+08	-1.2	-0.02	0.983
	A	4.80E+08	1.80E+08	3.9	0.65	0.526
Firm size, t=2006	B	40.349	132.77	-11	-0.21	0.836
	A	40.349	24.813	1.9	0.59	0.569
(Firm size)^2, t=2006	B	4565.3	1.40E+06	-8.4	-0.16	0.876
	A	4565.3	1904.6	0	0.71	0.490
Share young, t=2006	B	0.167	0.093	86.8	2.84	0.005
	A	0.167	0.145	26.5	0.47	0.645
(Share young)^2, t=2006	B	0.0365	0.013	70.1	2.91	0.004
	A	0.0365	0.026	31.2	0.58	0.573

Table 2 Descriptive statistics at the firm level for 2006-2010. Minimum of two employees

	Mean	Median	Std.dev.	Min	Max	Obs
Total wage sum (100 SEK)						
Other firms	71649.76	10057.16	995042.9	20.698	1.37E+08	728 587
Firms in control industries	37027.60	7109.313	419277.1	20.698	3.8E+08	29 707
Restaurant	11168.49	5195.198	64246.77	27.3	4035883	45 314
Hairdresser	6806.106	4708.463	8113.156	54.6	190697.9	8 859
Wage sum per employee (100 SEK)						
Other firms	2664.876	2559.46	1167.121	10.349	63698.77	728 587
Firms in control industries	1958.054	1903.698	783.964	10.349	11912.71	29 707
Restaurant	1511.264	1459.299	640.1976	13.65	16606.55	45 314
Hairdresser	1769.29	1767.287	667.6847	27.3	9362.324	8 859
No. of employees						
Other firms	23.759	4	327.694	2	46345	728 587
Firms in control industries	17.894	4	215.5722	2	20747	29 707
Restaurant	6.788	3	34.187	2	2238	45 314
Hairdresser	3.687	2	3.677	2	70	8 859
Share of ≤ 25 year old						
Other firms	0.119	0	0.188	0	1	728 587
Firms in control industries	0.160	0	0.211	0	1	29 707
Restaurant	0.266	0.25	0.279	0	1	45 314
Hairdresser	0.267	0.25	0.279	0	1	8 859

Note: Wage variables are adjusted for inflation using the base year of 2013. Outliers are excluded.

Figure 1 Wage development within restaurants, hairdressers and control firms in the pretreatment period. Minimum of two employees

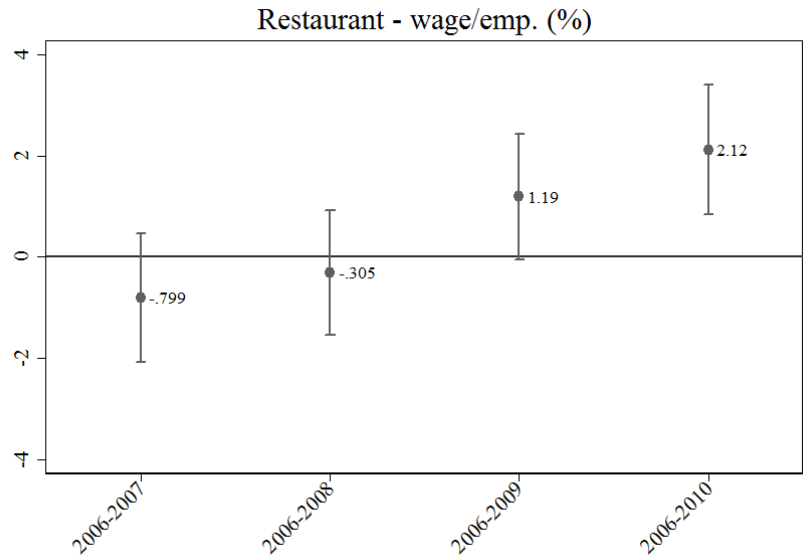


Note: Average ln(wage/employee) in the pretreatment years, including all firms within the treatment and control industries. Outliers are excluded.

Figure 2 Effects of staff registers on wages per employee. Restaurant and hairdresser firms with at least two employees

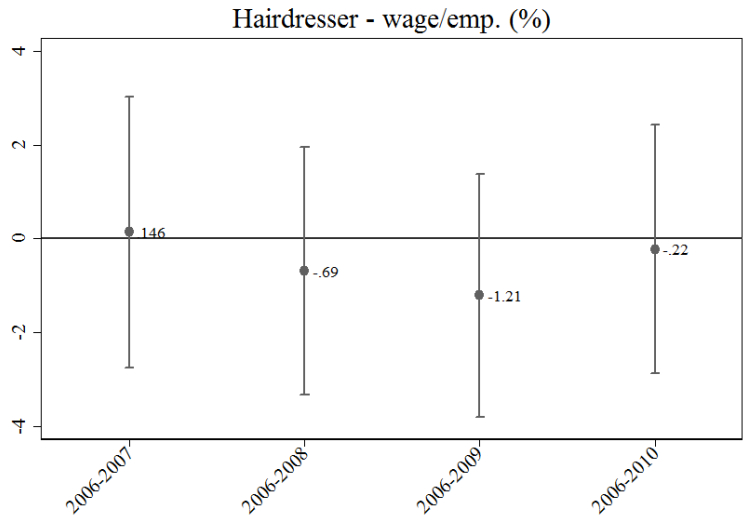
Note: Point estimates and their 95 % confidence intervals from the estimation of equation (2). Firm clustered standard errors. Only surviving firms remaining within the same industry are included. Outliers are excluded.

Figure 3 Effects of staff registers on wages per employee. Restaurant firms with at least two employees



Note: Point estimates and their 95 % confidence intervals from the estimation of equation (2). Firm clustered standard errors. Only surviving firms remaining within the same industry are included. Outliers are excluded.

Figure 4. Effects of staff registers on wages per employee. Hairdresser firms with at least two employees



Note: Point estimates and their 95 % confidence intervals from the estimation of equation (2). Firm clustered standard errors. Only surviving firms remaining within the same industry are included. Outliers are excluded

Appendix 1. Propensity Score Matching

Table A1 Probit model estimation. Propensity score matching.

Probit model estimation	
Wage/emp., t=2006	0.0070325
	(0.0044114)
(Wage/emp.)^2, t=2006	-2.29E-06**
	(1.14E-06)
ΔWage/emp., t=2005	0.0706564*
	(0.0372614)
(ΔWage/emp.)^2, t=2005	-0.000645**
	(0.0002733)
ΔWage/emp., t=2004	-0.0005989
	(0.0013106)
(ΔWage/emp.)^2, t=2004	1.26E-06
	(2.08E-06)
Size, t=2006	0.0000173
	(0.0000197)
(Size)^2, t=2006	1.41E-10
	(9.10E-11)
Firm size, t=2006	0.0308532***
	(0.0093043)
(Firm size)^2, t=2006	-0.0000903***
	(0.0000334)
Share young, t=2006	2.813979
	(8.432039)
(Share young)^2, t=2006	-8.394332
	(23.66927)
Constant	-8.992269**
	(4.274398)
Obs.	754
Pseudo R2	0.4826

Note: Dependent variable is equal to one for all five-digit restaurant and hairdresser industries in 2007 and equal to zero for all other five-digit industries in 2007. Standard errors clustered on the five-digit industry level. ***p<0.01, **p<0.05, *p<0.1

Table A2 Treated and control industries

Treated industries		Control industries	
SNI2002	**Description**	**SNI2002**	**Description**
55300	Restaurants	52121	Other retail sale in department stores and the like
		85328	Day-care activities for disabled persons
		74701	Cleaning of premises
		36630	Other manufacturing n.e.c.
		64120	Courier activities other than national post activities
55510	Canteens	36630	Other manufacturing n.e.c.
		52121	Other retail sale in department stores and the like
		85328	Day-care activities for disabled persons
		74701	Cleaning of premises
		64120	Courier activities other than national post activities
55522	Catering for hospitals	52472	Retail sale of newspapers and magazines
		01217	Farmers of animals, mized, mainly cattle
		74811	Portrait photography
		52483	Retail sale of watches and clocks
		63303	Tourist assistance
55529	Other catering	01301	Mixed farming, mainly crops and market garden produce
		52442	Retail sale of sugar confectionery
		05012	Other sea water fishing
		52632	Ambulatory and occasional retail sale of other goods
		01122	Growing of nursery products etc. in the open
93021	Hairdressing	93012	Washing and drycleaning for households
		01253	Bee keeping, raising of worms and other animals
		52279	Retail sale of food in specialized stores n.e.c.
		55102	Lodging activities of conference activities
		52410	Retail sale of textiles
55523	Catering for schools, welfare and other institutions	01302	Mixed farming, mainly animals
		15810	Manufacture of bread; manufacture of fresh pastry goods and cakes
		92310	Artistic and literatury creation and interpretation
		01137	Growers and crops and market garden produce, mixed, mainly fruit, berries, nuts etc.
		01113	Growing of potatoes
55521	Catering for the transport sector	36630	Other manufacturing n.e.c.
		52121	Other retail sale in department stores and the like
		85328	Day-care activities for disabled persons
		74701	Cleaning of premises
		64120	Courier activities other than national post activities

Note: Industry code 93021 constitutes the hairdresser sector. Other industry codes jointly constitute the restaurant sector.

Appendix 2. Estimation result tables

Table A3 Regression results. Restaurant and hairdresser firms with at least two employees. Estimates in Figure 2 correspond to the ATE estimate in the fourth column

2006-2007					**2006-2008**				
	1	2	3	4		1	2	3	4
TI	-0.226***	-0.126***	-0.126***	-0.232***	TI	-0.220***	-0.143***	-0.143***	-0.230***
	(0.00782)	(0.0349)	(0.0374)	(0.00766)		(0.00796)	(0.0286)	(0.0360)	(0.00785)
TP	0.0758***	0.0748***	0.0748***	0.0752***	TP	0.0820***	0.0809***	0.0809***	0.0800***
	(0.00888)	(0.00876)	(0.00541)	(0.00430)		(0.00778)	(0.00767)	(0.00501)	(0.00410)
ATE	-0.0176	-0.0155	-0.0155**	-0.00759	ATE	-0.0147	-0.0127	-0.0127*	-0.00489
	(0.0110)	(0.0109)	(0.00704)	(0.00587)		(0.00972)	(0.00957)	(0.00656)	(0.00561)
Constant	7.437***	7.469***	7.469***	7.422***	Constant	7.450***	7.495***	7.495***	7.431***
	(0.00630)	(0.0340)	(0.0360)	(0.00602)		(0.00637)	(0.0276)	(0.0346)	(0.00609)
Observations	27,948	27,948	27,948	27,948	Observations	37,082	37,082	37,082	37,082
R-squared	0.066	0.093	0.093		R-squared	0.070	0.099	0.099	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				15,535	Number of firms				14,137

2006-2009					**2006-2010**				
	1	2	3	4		1	2	3	4
TI	-0.218***	-0.147***	-0.147***	-0.223***	TI	-0.209***	-0.147***	-0.147***	-0.212***
	(0.00818)	(0.0247)	(0.0335)	(0.00796)		(0.00832)	(0.0221)	(0.0324)	(0.00815)
TP	0.0905***	0.0892***	0.0892***	0.0908***	TP	0.0992***	0.0980***	0.0980***	0.100***
	(0.00750)	(0.00738)	(0.00478)	(0.00411)		(0.00736)	(0.00724)	(0.00481)	(0.00421)
ATE	0.000240	0.00259	0.00259	0.00644	ATE	0.0119	0.0143	0.0143**	0.0151***
	(0.00941)	(0.00927)	(0.00639)	(0.00567)		(0.00928)	(0.00914)	(0.00646)	(0.00581)
Constant	7.465***	7.502***	7.502***	7.443***	Constant	7.471***	7.516***	7.516***	7.448***
	(0.00651)	(0.0237)	(0.0320)	(0.00609)		(0.00660)	(0.0211)	(0.0309)	(0.00619)
Observations	44,067	44,067	44,067	44,067	Observations	49,514	49,514	49,514	49,514
R-squared	0.069	0.098	0.098		R-squared	0.066	0.096	0.096	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				12,787	Number of firms				11,620

Dependent variable: ln(wage sum/employee). ATE*100 = point estimates in the paper. Standard errors within parentheses.
*p<0.10, **p<0.05, ***p<0.01

Table A4 Regression results. Restaurant firms with at least two employees. Estimates in Figure 3 correspond to the ATE estimate in the fourth column

2006-2007					**2006-2008**				
	1	2	3	4		1	2	3	4
TI	-0.270***	-0.145***	-0.145***	-0.273***	TI	-0.264***	-0.152***	-0.152***	-0.270***
	(0.00841)	(0.0407)	(0.0457)	(0.00828)		(0.00856)	(0.0336)	(0.0441)	(0.00851)
TP	0.0718***	0.0719***	0.0719***	0.0756***	TP	0.0789***	0.0787***	0.0787***	0.0807***
	(0.00942)	(0.00936)	(0.00580)	(0.00459)		(0.00824)	(0.00817)	(0.00541)	(0.00440)
ATE	-0.0140	-0.0141	-0.0141*	-0.00799	ATE	-0.00912	-0.00887	-0.00887	-0.00305
	(0.0118)	(0.0118)	(0.00771)	(0.00648)		(0.0104)	(0.0104)	(0.00723)	(0.00624)
Constant	7.451***	7.470***	7.470***	7.435***	Constant	7.463***	7.496***	7.496***	7.443***
	(0.00669)	(0.0340)	(0.0360)	(0.00642)		(0.00675)	(0.0276)	(0.0346)	(0.00650)
Observations	23,429	23,429	23,429	23,429	Observations	30,906	30,906	30,906	30,906
R-squared	0.090	0.103	0.103		R-squared	0.095	0.111	0.111	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				13,008	Number of firms				11,729

2006-2009					**2006-2010**				
	1	2	3	4		1	2	3	4
TI	-0.261***	-0.156***	-0.156***	-0.264***	TI	-0.255***	-0.170***	-0.170***	-0.255***
	(0.00881)	(0.0295)	(0.0432)	(0.00867)		(0.00900)	(0.0269)	(0.0441)	(0.00891)
TP	0.0884***	0.0877***	0.0877***	0.0910***	TP	0.0976***	0.0971***	0.0971***	0.101***
	(0.00793)	(0.00785)	(0.00516)	(0.00443)		(0.00779)	(0.00772)	(0.00520)	(0.00453)
ATE	0.00785	0.00858	0.00858	0.0119*	ATE	0.0204**	0.0209**	0.0209***	0.0212***
	(0.0101)	(0.0100)	(0.00707)	(0.00633)		(0.0100)	(0.00995)	(0.00718)	(0.00651)
Constant	7.477***	7.503***	7.503***	7.454***	Constant	7.483***	7.517***	7.517***	7.459***
	(0.00689)	(0.0237)	(0.0320)	(0.00651)		(0.00699)	(0.0212)	(0.0309)	(0.00660)
Observations	36,443	36,443	36,443	36,443	Observations	40,762	40,762	40,762	40,762
R-squared	0.094	0.112	0.112		R-squared	0.092	0.109	0.109	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				10,478	Number of firms				9,423

Dependent variable: ln(wage sum/employee). ATE*100 = point estimates in the paper. Standard errors within parentheses.
*p<0.10, **p<0.05, ***p<0.01

Table A5 Regression results. Hairdresser firms with at least two employees. Estimates in Figure 4 correspond to the ATE estimate in the fourth column.

2006-2007					**2006-2008**				
	1	2	3	4		1	2	3	4
TI	-0.000484	0.0889	0.0889	-0.0193	TI	-0.00533	0.0822	0.0822	-0.0279
	(0.0205)	(0.118)	(0.215)	(0.0200)		(0.0210)	(0.0937)	(0.205)	(0.0201)
TP	0.0970***	0.0947***	0.0947***	0.0671***	TP	0.0992***	0.0970***	0.0970***	0.0720***
	(0.0250)	(0.0249)	(0.0149)	(0.0126)		(0.0222)	(0.0220)	(0.0132)	(0.0113)
ATE	-0.0309	-0.0286	-0.0286	0.00146	ATE	-0.0373	-0.0351	-0.0351**	-0.00690
	(0.0294)	(0.0292)	(0.0177)	(0.0147)		(0.0259)	(0.0257)	(0.0159)	(0.0134)
Constant	7.340***	7.251***	7.251***	7.334***	Constant	7.361***	7.274***	7.274***	7.348***
	(0.0175)	(0.118)	(0.215)	(0.0168)		(0.0180)	(0.0931)	(0.205)	(0.0169)
Observations	4,519	4,519	4,519	4,519	Observations	6,176	6,176	6,176	6,176
R-squared	0.008	0.022	0.022		R-squared	0.008	0.022	0.022	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				2,527	Number of firms				2,408

2006-2009					**2006-2010**				
	1	2	3	4		1	2	3	4
TI	-0.0163	0.0624	0.0624	-0.0359*	TI	-0.00364	0.130*	0.130	-0.0272
	(0.0216)	(0.0807)	(0.192)	(0.0201)		(0.0216)	(0.0736)	(0.219)	(0.0206)
TP	0.101***	0.0993***	0.0993***	0.0869***	TP	0.107***	0.105***	0.105***	0.0911***
	(0.0215)	(0.0214)	(0.0124)	(0.0111)		(0.0209)	(0.0207)	(0.0126)	(0.0114)
ATE	-0.0264	-0.0250	-0.0250	-0.0121	ATE	-0.0153	-0.0133	-0.0133	-0.00220
	(0.0250)	(0.0249)	(0.0152)	(0.0132)		(0.0242)	(0.0240)	(0.0153)	(0.0135)
Constant	7.384***	7.305***	7.305***	7.364***	Constant	7.390***	7.256***	7.256***	7.370***
	(0.0186)	(0.0800)	(0.191)	(0.0168)		(0.0186)	(0.0728)	(0.219)	(0.0175)
Observations	7,624	7,624	7,624	7,624	Observations	8,752	8,752	8,752	8,752
R-squared	0.009	0.022	0.022		R-squared	0.010	0.026	0.026	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				2,309	Number of firms				2,197

Dependent variable: ln(wage sum/employee). ATE*100 = point estimates in the paper. Standard errors within parentheses.
*p<0.10, **p<0.05, ***p<0.01

Appendix 3. Robustness analysis

Table A6 Regression results. Restaurant firms – staying individuals. Firms having at least two employees. Baseline results are found in the fourth column

2006-2007					**2006-2008**				
	1	2	3	4		1	2	3	4
TI	-0.243***	-0.120*	-0.120	-0.257***	TI	-0.229***	-0.0457	-0.0457	-0.244***
	(0.00729)	(0.0693)	(0.103)	(0.00751)		(0.00788)	(0.0601)	(0.120)	(0.00833)
TP	0.105***	0.103***	0.103***	0.105***	TP	0.102***	0.100***	0.100***	0.104***
	(0.00681)	(0.00676)	(0.00414)	(0.00311)		(0.00624)	(0.00619)	(0.00396)	(0.00323)
ATE	0.00602	0.00736	0.00736	0.0217***	ATE	0.0263***	0.0276***	0.0276***	0.0383***
	(0.0103)	(0.0102)	(0.00714)	(0.00569)		(0.00962)	(0.00954)	(0.00690)	(0.00589)
Constant	7.536***	7.525***	7.525***	7.530***	Constant	7.597***	7.528***	7.528***	7.587***
	(0.00483)	(0.0651)	(0.0972)	(0.00442)		(0.00510)	(0.0562)	(0.115)	(0.00474)
Observations	32,368	32,368	32,368	32,368	Observations	40,998	40,998	40,998	40,998
R-squared	0.075	0.090	0.090		R-squared	0.063	0.079	0.079	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				17,775	Number of firms				15,290

2006-2009					**2006-2010**				
	1	2	3	4		1	2	3	4
TI	-0.222***	-0.131**	-0.131	-0.231***	TI	-0.226***	-0.0912*	-0.0912	-0.236***
	(0.00845)	(0.0533)	(0.0950)	(0.00877)		(0.00884)	(0.0506)	(0.103)	(0.00941)
TP	0.105***	0.104***	0.104***	0.108***	TP	0.113***	0.112***	0.112***	0.116***
	(0.00622)	(0.00617)	(0.00407)	(0.00347)		(0.00620)	(0.00616)	(0.00405)	(0.00356)
ATE	0.0439***	0.0451***	0.0451***	0.0503***	ATE	0.0614***	0.0624***	0.0624***	0.0655***
	(0.00974)	(0.00966)	(0.00717)	(0.00637)		(0.00987)	(0.00979)	(0.00747)	(0.00685)
Constant	7.637***	7.650***	7.650***	7.626***	Constant	7.661***	7.629***	7.629***	7.650***
	(0.00540)	(0.0491)	(0.0869)	(0.00488)		(0.00556)	(0.0464)	(0.0929)	(0.00497)
Observations	47,331	47,331	47,331	47,331	Observations	51,646	51,646	51,646	51,646
R-squared	0.054	0.069	0.069		R-squared	0.053	0.068	0.068	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				13,420	Number of firms				11,781

Dependent variable: ln(wage sum/employee). ATE*100 = point estimates in the paper. Standard errors within parentheses.
*p<0.10, **p<0.05, ***p<0.01

Table A7 Regression results. Hairdresser firms – staying individuals. Firms having at least two employees. Baseline results are found in the fourth column

2006-2007					**2006-2008**				
	1	2	3	4		1	2	3	4
TI	-0.0357**	-0.0917***	-0.0917***	-0.0505***	TI	-0.0340*	-0.103***	-0.103***	-0.0488***
	(0.0165)	(0.0238)	(0.0309)	(0.0166)		(0.0183)	(0.0244)	(0.0325)	(0.0184)
TP	0.0910***	0.0907***	0.0907***	0.0790***	TP	0.0920***	0.0912***	0.0912***	0.0769***
	(0.0165)	(0.0165)	(0.0102)	(0.00750)		(0.0160)	(0.0160)	(0.0103)	(0.00823)
ATE	0.00438	0.00462	0.00462	0.0190*	ATE	-0.00848	-0.00768	-0.00768	0.0115
	(0.0235)	(0.0234)	(0.0155)	(0.0112)		(0.0224)	(0.0224)	(0.0155)	(0.0125)
Constant	7.412***	7.468***	7.468***	7.406***	Constant	7.445***	7.514***	7.514***	7.434***
	(0.0116)	(0.0207)	(0.0283)	(0.0111)		(0.0130)	(0.0208)	(0.0297)	(0.0127)
Observations	6,583	6,583	6,583	6,583	Observations	8,724	8,724	8,724	8,724
R-squared	0.011	0.015	0.015		R-squared	0.009	0.014	0.014	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				3,660	Number of firms				3,331

2006-2009					**2006-2010**				
	1	2	3	4		1	2	3	4
TI	-0.0321	-0.109***	-0.109***	-0.0488**	TI	-0.00800	-0.0989***	-0.0989***	-0.0192
	(0.0196)	(0.0245)	(0.0332)	(0.0196)		(0.0199)	(0.0239)	(0.0332)	(0.0200)
TP	0.111***	0.110***	0.110***	0.0942***	TP	0.117***	0.116***	0.116***	0.107***
	(0.0163)	(0.0162)	(0.0111)	(0.00936)		(0.0161)	(0.0160)	(0.0114)	(0.00994)
ATE	-0.0271	-0.0262	-0.0262	-0.00770	ATE	-0.0189	-0.0172	-0.0172	-0.0106
	(0.0226)	(0.0226)	(0.0164)	(0.0135)		(0.0223)	(0.0222)	(0.0164)	(0.0139)
Constant	7.478***	7.554***	7.554***	7.466***	Constant	7.493***	7.584***	7.584***	7.478***
	(0.0141)	(0.0204)	(0.0302)	(0.0137)		(0.0144)	(0.0196)	(0.0303)	(0.0144)
Observations	10,530	10,530	10,530	10,530	Observations	11,853	11,853	11,853	11,853
R-squared	0.010	0.015	0.015		R-squared	0.008	0.016	0.016	
Industry FE	No	Yes	Yes	No	Industry FE	No	Yes	Yes	No
Firm RE	No	No	No	Yes	Firm RE	No	No	No	Yes
Firm clustered s.e's	No	No	Yes	Yes	Firm clustered s.e's	No	No	Yes	Yes
Number of firms				3,104	Number of firms				2,875

Dependent variable: ln(wage sum/employee). ATE*100 = point estimates in the paper. Standard errors within parentheses.
*p<0.10, **p<0.05, ***p<0.01

Table A8 Staying individuals. Balancing test for matched treated and control industries

Variables	B/A = Before/After matching	Mean			T-test	
		Treated	Control	Bias (%)	t	p-value
Wage/emp.., t=2006 (stay 05-06)	B	2058.7	2944	-145.7	-2.91	0.004
	A	2058.7	2009.4	8.1	0.34	0.743
(Wage/emp.)^2, t=2006 (stay 05-06)	B	4.30E+06	9.30E+06	-127.8	-2.45	0.015
	A	4.30E+06	4.10E+06	6	0.39	0.702
ΔWage/emp., t=2005 (stay 04-05)	B	105.92	145.79	-30	-0.66	0.511
	A	105.92	116.95	-8.3	-0.24	0.815
(ΔWage/emp.)^2, t=2005 (stay 04-05)	B	19473	46684	-21.3	-0.4	0.687
	A	19473	18141	1	0.11	0.913
ΔWage/emp., t=2004 (stay 03-04)	B	49	145.82	-21.5	-0.43	0.669
	A	49	88.243	-8.7	-0.47	0.649
(ΔWage/emp.)^2, t=2004 (stay 03-04)	B	41234	3.80E+05	-5.2	-0.1	0.922
	A	41234	11446	0.5	1.23	0.241
Size, t=2006 (stay 05-06)	B	6939.1	3895.4	18.9	0.41	0.682
	A	6939.1	8322.7	-8.6	-0.22	0.832
(Size)^2, t=2006 (stay 05-06)	B	1.60E+08	4.00E+08	-4.4	-0.08	0.935
	A	1.60E+08	2.00E+08	-0.6	-0.15	0.884
Firm size, t=2006 (stay 05-06)	B	53.44	112.72	-8.4	-0.16	0.874
	A	53.44	21.717	4.5	0.85	0.411
(Firm size)^2, t=2006 (stay 05-06)	B	9907.7	9.90E+05	-7.4	-0.14	0.89
	A	9907.7	1730.1	0.1	0.92	0.375

Note: Variables are built upon individuals staying at least for two consecutive years at the same firm. For instance, 'stay 05-06' refers to individuals remaining at the same firm in years 2005-2006.

Table A9 Staying individuals. Matched control industries

Treated industries		Control industries	
SNI2002	**Description**	**SNI2002**	**Description**
55300	Restaurants	52112	Retail sale in other non-specialized stores with food, beverages and tobacco predominating
		64120	Courier activities other than national post activities
		74701	Cleaning of premises
		60211	Urban and suburban scheduled passenger transport
		85328	Day-care activities for disabled persons
55510	Canteens	52121	Other retail sale in department stores and the like
		52112	Retail sale in other non-specialized stores with food, beverages and tobacco predominating
		64120	Courier activities other than national post activities
		74701	Cleaning of premises
		60211	Urban and suburban scheduled passenger transport
55522	Catering for hospitals	01228	Breeding of horses etc.
		01124	Growing of flowers and ornamental plants under glass
		60220	Taxi operation
		93012	Washing and drycleaning for households
		01302	Mixed farming, mainly animals
55529	Other catering	52431	Retail sale of footwear
		52443	Retail sale of glassware, china and kitchenware
		01129	Growing of mushrooms etc.
		52260	Retail sale of tobacco products
		52486	Retail sale of games and toys
93021	Hairdressing	05025	Growing of aquatic plants
		52423	Retail sale of women's clothing
		17120	Preparation and spinning of woollen-type fibres
		52487	Retail sale of flowers and other plants
		01300	Growing of crops combined with farming of animals (mixed farming)
55523	Catering for schools, welfare and other institutions	52410	Retail sale of textiles
		01500	Hunting, trapping and game propagation including related service abilities
		01259	Raising and breeding of other animals
		01232	Raising of swine for slaughter
		01121	Growing of vegetables in the open
55521	Catering for the transport sector	52121	Other retail sale in department stores and the like
		52112	Retail sale in other non-specialized stores with food, beverages and tobacco predominating
		64120	Courier activities other than national post activities
		74701	Cleaning of premises
		36630	Other manufacturing n.e.c.

Note: Industry code 93021 constitutes the hairdresser sector. Other industry codes jointly constitute the restaurant sector.