



DALARNA
UNIVERSITY

Working papers in transport, tourism, information technology and microdata analysis

Firm relocation and firm profits: Evidence from the Swedish wholesale trade sector

Author: Zuzana Macuchova

Nr: 2015:06

Editor: Hasan Fleyeh

Working papers in transport, tourism, information technology and microdata analysis

ISSN: 1650-5581

© Authors

Firm relocation and firm profits: Evidence from the Swedish wholesale trade sector

Zuzana Macuchova

Dalarna University, SE- 791 88 Falun, Sweden.

Abstract: This study analyses the effects of firm relocation on firm profits, using longitudinal data on Swedish limited liability firms and employing a difference-in-difference propensity score matching method in the empirical analysis. Using propensity score matching, the pre-relocalization differences between relocating and non-relocating firms are balanced. In addition to that, a difference-in-difference estimator is employed in order to control for all time-invariant unobserved heterogeneity among firms. For matching, nearest-neighbour matching, using the one, two, and three nearest neighbours is employed. The balancing results indicate that matching achieves a good balance, and that similar relocating and non-relocating firms are being compared. The estimated average treatment on the treated effect indicates that relocation has a significant effect on the profits of the relocating firms. In other words, firms that relocate increase their profits significantly, in comparison to what the profits would be had the firms not relocated. This effect is estimated to vary between 3 to 11 percentage points, depending on the length of the analysed period after relocation.

Keywords: Firm relocation, propensity score matching, longitudinal data, wholesale trade sector

JEL codes: L25, R12, L81

1 Introduction

It is generally acknowledged that there are three main components in the economic dynamics of a region: entry and exit of firms, growth and decline of existing enterprises, and the spatial redistribution of existing firms. While interest in new firm formation has a long tradition in regional science research, resulting in an extensive body of literature analysing the patterns, effects or factors influencing entry of new firms, the issue of firm relocation is studied more sparsely. First in the recent decade, there emerged several studies on various aspects of firm relocation, predominately in countries where detailed firm-level databases are available. Thus, the empirical evidence on relocation of firms covers mainly the Netherlands (van Dijk and Pellenburg, 2000; Brouwer et al., 2004; Stam, 2007; Knobens et al., 2008; Knobens, 2011; Wetering and Knobens, 2013; Sleutjes and Völker, 2012; Kronenberg, 2013), Portugal (Holl, 2004), Switzerland (Bodenmann and Axhausen, 2010) and Sweden (Daunfeldt et al., 2013a, Håkansson et al., 2013).

An important strain of firm relocation research focuses on factors influencing a firm's decision to relocate. There is nowadays broad consensus in this literature, that a firm's relocation decision is mainly driven by firm-internal factors (Hayter, 1997). In particular,

firm age and firm size are repeatedly reported to be important factors ‘pushing’ firms from their present location (van Dijk and Pellenbarg, 2000; Stam, 2007; Knoben, 2011; Håkansson et al., 2013). Besides firm-internal factors, numerous studies also consider spatial characteristics of the firm’s current location on the decision to relocate. The empirical evidence on these factors is less conclusive (Knoben, 2011; Wetering and Knoben, 2013; Sleutjes and Völker, 2012; Kroneberg, 2013).

While the findings from these studies are very insightful in explaining factors affecting the firm’s relocation decision, few studies examine another aspect of firm relocation, namely, the effects that relocation has on the firms that relocate. Nakosteen and Zimmer (1987) and Knoben et al. (2008) are two exceptions, examining the post-relocation outcomes of relocating firms. Nakosteen and Zimmer (1987) examined changes in sales and employment growth for US manufacturing firms, concluding that migrant firms experienced higher rates of employment growth, while no such effect was found on sales growth. In a more recent study, Knoben et al. (2008) analysed, among other factors, also the effects of firm relocation on profits and R&D performance among Dutch automation industry firms. The authors reported that relocating firms experience a decrease in their profitability in the two years immediately after relocation, while profitability increases five years after firm relocation.

The lack of studies on the effects of relocation in regional science literature is surprising, for at least two reasons. Firstly, to investigate the factors that influence a firm’s decision to relocate is one important part of the knowledge, but at least as important is knowing what happens to firms after relocation. Drawing a parallel to human migration studies, a great amount of interest in this field is focused upon the outcomes of human relocation. For example, the effects of migration on individuals’ earnings, or on labour market outcomes, are some of the questions under scrutiny in this literature. Secondly, part of the empirical studies searching for firm relocation determinants rely on the neoclassical approach, which assumes firms to be profit maximizers, with the relocation itself being driven by firms’ expected earnings. Meanwhile, with the exception of Nakosteen and Zimmer (1987), and Knoben et al. (2008), none of these studies evaluate if relocating firms actually experience higher earnings after relocation, compared to what the earnings would be, had not the firms relocated.

This paper aims to extend the literature on firm relocation by analysing the post-relocation profits of relocating firms. To this end, longitudinal data on limited liability firms in Sweden covering the period 1998-2008 is used, and return on total assets (ROA) is the measure of the firms’ profits adopted in this study. The longitudinal nature of the data allows the identification of the relocation patterns of each relocating firm, and besides that, the data provides a rich set of characteristics on both relocating and non-relocating firms. In the paper, propensity score matching, in combination with a difference-in-difference estimator is used, in order to estimate the average treatment effect on the treated (ATT) of relocation, i.e. the causal effect of relocation on the profitability of firms is under scrutiny. The propensity score matching accounts for the fact that relocating firms are expected to be a non-random sample of firms (e.g. Håkansson et al., 2013, Knoben and

Oerlemans, 2008). Furthermore, by using a difference-in-difference estimator, the study controls for the time-invariant differences between relocating and non-relocating firms, i.e. the possible selection on unobservables, an issue frequently reported from studies evaluating post-migration effects of human migration (see e.g. Eliasson et al., 2012). The key results indicate that relocating firms experience higher profits in the post-relocation period, compared to what the profits would be without relocation. These results are robust to a number of different specifications of the matching estimator. To the author's best knowledge, this is the first time a longitudinal firm-level dataset and a difference-in-difference matching estimator are combined, in order to analyse the effects of firm relocation on the return on assets of relocating firms. Even though Nakosteen and Zimmer's (1987) theoretical model allows for testing the after-relocation profits of relocating firms, the nature of their dataset¹ hampered testing this hypothesis empirically, and, in addition to that, firm profits are in their study proxied with employment and sales growth. Contrary to Nakosteen and Zimmer, Knoben et al. (2008) had access to firm profit measures, however, the nature of the data collection process in this study did not allow controlling for potential sample-selection bias.²

The paper begins with a presentation of a firm relocation model and a review of the determinants of firm relocation, both presented in Section 2. In Section 3, the identification strategy used in this paper is outlined, and in Section 4, the data and measurements are presented. In Section 5, the empirical results are presented, while Section 6 concludes the paper.

2 Theoretical framework

2.1. Modelling firm relocation

The primary goal of this paper is to empirically measure the effects of relocation on the earnings of relocating firms. However, to give some theoretical background, the paper starts by presenting the manner in which previous studies have theoretically modelled a firms' decision to relocate. To this end, a theoretical framework similar to the one put forward by Nakosteen and Zimmer (1987), and subsequently adopted in other relocation studies (e.g. van Dijk and Pellenbarg 2000, Håkansson et al., 2013) is presented in this section.

Following Nakosteen and Zimmer (1987), a firm's decision to relocate is assumed to arise from the profit maximizing behaviour of the firm. Furthermore, firms are assumed to

1 Naakosteen and Zimmer's (1987) study employs Dun and Bradstreet Corporation records, a repeated cross-sectional data covering firms, applying for credits, for the years 1970 and 1980. This implies a possible endogeneity issue, when both the relocation event and the outcome of the relocation are measured in the same time period.

2 Knoben et al. (2008) use a data set of 203 firms, collected through a questionnaire survey among managing directors of 2,553 firms. Even though the authors perform a non-response analysis, the issue of the self-selection of firms into the sample cannot be ruled out.

be price takers both in product and factor markets. The expected profits E_{ijm} of firm i , located in region m , and active in industry j can be expressed as:

$$E_{ijr} = E(X_i, K_j, R_m) \quad (1)$$

where X_i represents firm-specific factors, K_j denotes factors related to the industry j , in which firm i operates, and where R_m captures non-industry specific variables of the region where the firm is located.

The firm's decision to relocate is then viewed as a capital investment project, where firm i 's discounted expected net gain in profits if relocating from origin m to destination s and incorporating the moving costs, C , is given by:

$$V_{i,m,s} = \int_t^T (E_{im} - E_{is})e^{-\tau t} dt - C_{r,s} \quad (2)$$

In Eq. (2), E_{im} and E_{is} represent the expected profits in regions m and s during period t , respectively; $C_{r,s}$ is the one-time cost of moving between the two regions, T is the investment horizon, and τ is the rate at which profits and costs are discounted. Firm i relocates if the expected net gain from relocation, $V_{i,m,s}$, is positive.

Most previous studies relying on the neoclassical approach to relocation have assumed profit maximization in accordance with Eq. (2) as the reason for relocation, and then moved on to investigate the determinants of firm relocation (see e.g. Nakosteen and Zimmer, 1987). This paper instead questions and investigates the validity of the predictions from theoretical models used in previous studies, that of relocating firms aiming for and achieving profit increases. For if firms use Eq. (2) as their decision rule and are able to correctly estimate E_{ir} and E_{is} , relocating firms should increase their profits due to the relocation, all else being equal.

In order to analyse this hypothesis, this paper employs a rich firm-level data, and somewhat different empirical approach, compared to the original paper of Nakosteen and Zimmer (1987). In the first step of the analysis, non-relocating firms having a similar probability of relocation as the relocating firms based on observables are identified. Then, in a second step of the empirical analysis, a difference-in-difference estimator controlling for time-invariant heterogeneity between relocating and non-relocating firms is used to assess the effects of relocation on relocating firms.

2.2. Firm relocation determinants

Firm-related factors are assumed to be the main drivers of firm's relocation decision (Hayter, 1997). There is a relatively clear consensus in the literature that among these, firm age and firm size are particularly correlated with firm relocation – nearly every empirical study, regardless of industrial branch or geographical region under scrutiny, reports that younger and smaller firms are more prone to relocate, compared to older and

larger firms (see e.g. Stam, 2007; Knoblen, 2011; Wetering and Knoblen, 2013; Håkansson et al., 2013). A possible explanation put forward by Garnsey (1998) is that older firms have lower relocation propensity because they have reached a stage of stabilization and stop growing. Furthermore, as reported in Stam (2007), Knoblen et al. (2008) and Knoblen et al. (2011) older firms are often more embedded into their environment through employees and network relationships and this makes them less prone to relocate. The territorial embeddness and relocation costs might also explain why smaller firms move more easily, compared to larger firms (Knoblen and Oerlemans, 2008; Brouwer et al., 2004). Another factor reported by several authors as a reason for firm relocation is firm growth. De Bok and van Ort (2011) point out that firms experiencing growth in the number of employees have a higher probability to relocate to other premises, and Håkansson et al. (2013) also found a positive correlation between growth in sales and the firm's propensity of relocate. However, as reported by Brouwer et al. (2004), also firms experiencing a decrease in their size are more prone to relocate.

Besides firm-related factors, recent studies also acknowledge the importance of external factors for the firm's relocation decision. Increasing attention in firm relocation literature is paid to the characteristics of the geographical location, where the firm operates (see Knoblen et al., 2008; Knoblen and Oerlemans, 2008; de Bok and van Oort, 2011; Håkansson et al., 2013). In this context, the authors usually distinguish between factors more closely related to the industry structure, in the geographical location of the firm, and non-industry related factors.

Among the factors related to the industry structure, two kinds of spatial externalities are commonly mentioned – localization and urbanization economies. Localization economies refer to the effects caused by having a large number of firms operating in the same industrial branch located in geographical proximity to each other, whereas urbanization economies emphasize the effects of proximity of firms from different industries closely co-located in one place. The general idea behind both localization and urbanization economies is that the spatial concentration of firms, both in the same industrial branch and in related branches, produces spatial externalities that the firm does not have any direct costs for. These externalities are in general expected to bring cost-saving benefits and productivity gains to firms localized in the region. Firms located in proximity to other firms in the same sector can take advantage from potential knowledge-spillover effects, or from the existence of nearby specialized suppliers (Figueiredo et al.; 2002, Holl, 2004; Arauzo-Carod, 2009). In the context of firm relocation studies, localization and agglomeration economies are in general expected to work as 'keep' factors, e.g. firms located in regions characterized by localization and agglomeration economies are expected to be less likely to relocate (de Bok and van Oort, 2011; Knoblen, 2011). However, while there is a clear theoretical consensus that the spatial concentrations of firms play a role, the empirical evidence in the literature is somewhat mixed. A positive effect of spatial concentration is reported in Wetering and Knoblen (2013), where a high level of specialization and urbanization functions as keep factors, i.e. firms located in regions characterized with higher spatial concentration of firms active in the same industry are less likely to move to other regions. This result the authors explain by the availability of specialized labour and other inputs, and low transportation costs. On the other hand,

Knoben (2011) points out that urbanization economies lure firms out from their own municipality, to search for another location, while Kronenberg (2013) did not find any effect of urbanization economies on the relocation of firms. Apparently, these results indicate that not all kinds of concentration of economic activity are necessarily beneficial for the firm. There could be agglomeration diseconomies related to the higher density of firms as well, e.g. increased competition or an increased demand for inputs driving up the costs of such inputs.

Non-industry location-related factors relate to the geographical environment where the firm is located. However, instead of addressing the spatial externalities generated by the co-location of firms in the same industrial branch, or by the presence of other firms in general, these factors describe other characteristics of the location. The educational level in the region is often argued to affect the entrepreneurial activity in general, and the supply of skilled labour in particular. The knowledge spillover theory of entrepreneurship (Acs et al., 2009), implies that entrepreneurship is more successful in regions with access to an educated labour force, as more highly educated people are expected to acquire more knowledge, valuable for firm operations. Empirical results on the effect of the educational level in a region are somewhat puzzling when compared to the theoretical findings. Wetering and Knoben (2013) control for the presence of a highly educated labour force, expecting that a lack of skilled labour might trigger firms to relocate from the region. However, their empirical results, on the contrary, show a higher level of firm relocation in regions characterized by a higher educational level. In another study, Nakosteen and Zimmer (1987) do not measure directly the level of higher education in the region, instead, education spending is used as a proxy variable. However, their findings are in line with Wetering and Knoben (2013), and – contrary to the expected hypothesis of a negative relationship between relocation from the region and high education spending – the authors report a positive relationship between educational spending and firm relocation. A possible explanation that Nakosteen and Zimmer (1987) suggest is that in the US, counties with high educational spending also have high corporate taxation, and that the latter drives firms out of the counties with higher spending on schooling. Finally, the political and institutional setting might also affect entrepreneurial activity, as attested in Baumol (1990). In previous studies on firm relocation in Sweden, (Håkansson et al., 2013; Daunfeldt et al., 2013a), the authors report a positive correlation between the firm’s out-relocation and left-of-centre-governed municipalities.

3 Empirical strategy

3.1. Matching estimator

In order to evaluate the effects of relocation on relocating firms, this paper employs difference-in-difference propensity score matching. Matching is one of the techniques within evaluation literature that enables the constructing of the missing potential outcome needed to evaluate effects of a treatment, as the propensity scores restrict matches to

suitable pairs of comparison units only. Since the seminal work of Rosenbaum and Rubin (1983), propensity score matching has become increasingly popular, mostly in labour economics (Heckman et. Al., 1998), but recently also in regional economics (i.e. Wenz, 2008; Bernini and Pellegrini, 2011; Wagner, 2011; di Cintio and Grassi, 2013; Kauder, 2014).

To justify the implementation of a matching strategy in this paper, consider the evaluation problem: I seek to investigate whether a relocating firm increases its profits, in the period after relocation, compared to what the profits would be had the firm not relocated. To correctly identify the effect of firm relocation on the profits of relocating firms this would require observations of firm profits, Y_i , of each relocating firm i in the sample in two potential states of the world: In state 1, given that firm i relocated, Y_i^1 , and in state 0, given that firm i did not relocate, Y_i^0 . The average effect of firm relocation on the profits of relocating firms would then be identified (after dropping the i subscript) as the ‘average treatment effect on the treated’, ATT:

$$\begin{aligned} ATT &= E(Y^1 - Y^0 | Relocation = 1) = \\ &= E(Y^1 | Relocation = 1) - E(Y^0 | Relocation = 1) \end{aligned} \quad (3)$$

The first term on the right-hand side of Eq. (3) is known, since the profits of relocating firms after the firms relocated are observed in the data. However, I do not observe the profits the relocating firms would have, had they not relocated, the second (counterfactual) term on the right-hand side of Eq. (3). Each firm i in the sample can either relocate or stay in the original location, it cannot be in state 0 and state 1 simultaneously and, thus, the second term on the right-hand side of Eq. (3) is inherently missing.

In this setting, *matching* is used to construct the counterfactual, missing outcome from Eq. (3), i.e. the profit of relocating firms, had they not relocated. Matching enables, under certain assumptions discussed further below, the use of information from a pool of non-relocating firms (control units), to identify what would happen to relocating firms (treated units), in the absence of the relocation (treatment). However, simply using all non-relocating firms to construct the profits of relocating firms in absence of the relocation, need not be a good estimation strategy. As shown in Imbens (2004), in non-experimental settings, treatment and control units may differ systematically, not only in their treatment statuses, but also in other *ex-ante* characteristics, and these characteristics might affect both assignment to the treatment and the outcomes of the treatment. In the case of firm relocation, findings from previous studies (Håkansson et al., 2013; Daunfeldt et al., 2013a) show that firm relocation is not a randomly driven event, instead, for example, younger and smaller firms are over-represented among relocating firms, while older and larger firms are less prone to change their geographical location. These results suggest considering *ex-ante* characteristics of firms in the evaluation problem as well.

Therefore *propensity score matching* is employed to restrict the control group to only those non-relocating firms having similar characteristics to the relocating firms in the period prior to relocation. Propensity scores, $e(X)$, introduced by Rosenbaum and Rubin

(1983), simplify significantly the process of finding suitable matches, since propensity scores summarize the observable *ex ante* characteristics X for each observation into one scalar variable only – the probability of being treated given the pre-treatment characteristics, that is: $e(X) = Pr(T=1|X) = Pr(T|X)$. The matching then comprises a matching on one value, the propensity score, rather than on several conditioning characteristics in X .

The propensity scores are balancing scores, which means that at each value of the propensity scores, the distribution of the baseline characteristics, grouped in the vector X , is the same in the treated and control group. In other words, the propensity score matching accounts for the observable *ex ante* differences among relocating and non-relocating firms. However, even after conditioning on observable factors, there may still be systematic differences between the treatment and control group, albeit unobservable. Selection on unobservables is a well-known phenomenon from studies on human migration and earnings.

As pointed out in, for example, Eliasson et al. (2012), some individuals may possess latent traits that make them stronger earners than otherwise comparable individuals who lack those traits. If those attributes also induce individuals to migrate, than migration is a partly self-selected process and the outcomes observed as consequence of migration, such as post-migration earnings, might reflect this selection on unobservables. Similarly to the process of human migration, in the case of firm relocation, there may be a risk that firms can possess different firm-internal factors, factors as, for example, quality of managerial staff, or access to human capital resources, which are difficult to capture in conventionally available data, and these latent factors might contribute both to the relocation decision of firms and the after-relocation profits. As a result, the after-relocation outcomes attributable in normal cases to the effect of the relocation might actually be influenced by unobservable characteristics of firms, rather than by relocation.

Thus, the propensity score matching is combined with a *difference-in-difference estimator* (DID) to overcome the potential selection bias caused by unobservable differences between relocating and non-relocating firms. As shown in Ham et al. (2011), the advantage of the DID estimator is that it, under certain assumptions, allows for time-invariant differences between treatment and control units.³ The implementation of DID estimator means technically that the main outcome variable of interest in the empirical analysis needs to be defined as prior/after change in profits, instead of levels (profits).

Thus, coming back to the example of the relocating firm i and the estimation of ATT outlined in Eq. (3). In state 1, given that firm i has relocated, the outcome variable is defined as: $Y_i^1 = E_{i1t} - E_{i0t'}$, while in state 0, given firm i did not relocate as: $Y_i^0 = E_{i0t} - E_{i0t'}$. In both cases, E_i represents the profits of firm i , while t and t' represents the

³ Ham et al. (2011) show that there is a trade-off related to the choice between choosing a cross-section (CS) and difference-in-difference estimator (DID). The DID matching estimator relies on an assumption of additive separability of error terms, while this assumption is not required by the CS estimator. On the other hand, the DID only needs to the CIA assumption, discussed in section 3.2, to hold after unobserved time invariant (separable) components that affect both relocation and profits, have been differenced out, while the CS estimator is required to hold without removing such time-invariant (separable) components. Given the settings of the evaluation problem, the DID estimator was chosen.

time period after relocation and prior to relocation respectively. The specification of the final difference-in-difference estimator is the following:

$$ATT = E[(Y_i^1 = E_{i1t} - E_{i0t'}) - (Y_i^0 = E_{i0t} - E_{i0t'}) | Relocation = 1] - E[(Y_i^1 = E_{i1t} - E_{i0t'}) | Relocation = 1] - E[(Y_i^0 = E_{i0t} - E_{i0t'}) | Relocation = 1] \quad (4)$$

3.2. Matching estimator assumptions and their plausibility

As outlined in the previous section, the difference-in-difference propensity score matching estimator under certain assumptions allows using profits of non-relocating firms to construct the missing counterfactual outcome from Eq. (4), the profits of relocating firms in the absence of relocation.

The first assumption required for the validity of the matching estimator, is a conditional independence assumption (CIA, see Rosenbaum and Rubin, 1983). Using the terminology of evaluation literature, this assumption means that after controlling for the observable differences between treatment and control units, the outcome of the treatment is independent of the treatment: $Y^0 \perp T / X$ with T as treatment, Y^0 is the outcome of the treatment unit, in the absence of the treatment, and X as the *ex-ante* characteristics. It can be shown (see Imbens, 2004) that since the propensity scores, $e(X)$, keep the balancing properties, if the conditional independence assumptions hold conditional on X , it will also hold conditional on the propensity score: $Y^0 \perp T / e(X)$. The implication of the conditional independence assumption, in this case, is that the expectation of the change in profits of non-relocating firms, conditional on the propensity score $e(X)$, is the same as the expectation for a relocating firm, conditional on $e(X)$, in the absence of relocation:⁴

$$[E(E_{i0t} - E_{i0t'} | e(X), Relocate_t = 1)] = E[E_{i0t} - E_{i0t'} | e(X), Relocate_t = 0]. \quad (5)$$

In other words, under CIA the term on the right-hand side of Eq. (5) provides the necessary counterfactual outcome, the profit change on relocating firms in the absence of the relocation, required to measure the effects of the relocation in the Eq. (4)

We cannot test the conditional independence assumption formally, however, as shown in Heckman et al. (1998) or Rubin and Thomas (1996), to increase the probability of this assumption to hold all variables which affect participation in treatment, and affect the outcome in the non-participation status, should be included in the matching procedure. Moreover, as shown by Imbens (2004), matching estimators have proven to be more effective when a large number of control units is available. I believe that selection of the conditioning variables, based on the results of recent firm relocation research, in

⁴ In Eq. (6), the outcome variable is stated in terms of before/after change of profits, instead of levels, since this study uses a DID matching estimator. DID weakens the identifying assumption for matching by allowing non observable time-invariant variables to influence performance (Bryson et al., 2002, in Bernini and Pellegrini, 2011).

combination with the rich dataset⁵ used providing a large control ‘reservoir’, makes it probable that the conditional independence assumption holds in the setting of this paper.

The second assumption required for a valid matching estimator is a common support condition. This assumption implies that there needs to be a positive probability of receiving treatment for all values of the conditioning variables in \mathbf{X} :

$$0 < P(\text{Relocation} = 1|\mathbf{X}) < 1. \quad (6)$$

In this case, the common support condition requires overlapping regions in the data, where non-relocating firms need to be nearby in terms of propensity scores distance to relocating firms, and furthermore, for each relocating firm, there should be reasonable number of non-relocating firms. The common support condition makes the difference between ordinary least squares (OLS) and matching. In an OLS framework, the identification of the effects would be based on all observations, also observations outside the common support, which might give rise to bias. In matching, observations being outside the common support are disregarded.

The common support condition is supported in the analysis by dropping the relocating firms (treatment group observations) whose propensity score is higher than the maximum, or lower than the minimum propensity score of the non-relocating firms (control group observations). In addition to that, following Heckman et al. (1998), 3% of the relocating firms for which the propensity score density of control observations is lowest, are dropped. In other words, in the second step of the matching procedure, the 3% of relocating firms, having least matches among non-relocating firms, are not included. To test the sensitivity of the matching estimator to the trimming level of 3%, trimming levels of 5% and 7%, respectively, were considered, as well and this did not affect the results of this study in any major way.

The third assumption that needs to be fulfilled for a valid matching estimator is a Stable Unit Treatment Value Assumption (SUTVA). In its general form, SUTVA requires that the outcome of one unit is not affected by treatment assignment on other unit. For the study, this assumption implies that a relocation of firm i from a region m does not influence the relocation decision of other firms, located in the same region. The SUTVA assumption might be violated in these settings. The relocation of one firm from a region might affect other firms operating in that region, and some of the non-relocating firms that are affected may appear in the control group. At the same time, the possible bias if the SUTVA assumption is violated should be small, since the number of relocating firms is small, and the probability that enough firms in the sample used to construct the control group is so severely affected by the relocating firms that it affects the results in any major way should be rather low.

⁵ The data used in this paper is discussed in Section 4.

4 Data and model specification

4.1. Data description

This study relies on a firm-level dataset of Swedish limited liability firms. In Sweden, all limited liability firms are obliged to submit their annual financial reports to the Swedish Patent and Registration Office. The data is compiled by PAR, a private consultancy firm and used for research purposes (e.g. Håkansson et al., 2013, and 2014; Daunfeldt et al., 2013a, and 2013b) and by decision-makers in Swedish commercial life. The financial reports data is ideally suited for studying the effects of firm relocation since the data contains, along with a variety of financial measures of firms, also information on the location of each firm. The longitudinal nature of the data allows identification of firm' relocation, and moreover, the longitudinal aspects of the data makes it possible to track all firms over time. In addition, since financial reports provide a relatively rich set of conditioning variables, including information on profits, costs, number of employees, value added and other variables included in firms' balance sheets, the richness of the financial reports data makes propensity score matching an appealing strategy to solve the fundamental evaluation problem.

The analysis is focused upon firms operating in the wholesale trade industry. The reason behind choosing this industry is a higher relocation frequency reported for firms within the wholesale trade, in comparison to firms in other sectors of the economy (see e.g. Knoben and Orleans, 2008; de Bok and van Oort, 2011; Sleutjes and Völker, 2012). The focus is also limited to single-plant firms, although the original dataset covers single-plant firms and single establishments of multiplant firms. The reason behind this selection is two-fold: firstly, single-plant firms adopt different strategies in how to cope with organizational changes, compared to multiplant firms (see e.g. Nakosteen and Zimmer, 1987; Pellenbarg et al., 2002) and, secondly, this decision is partly data-driven, since the nature of the dataset does not allow disintegration of the data to single-sites of multiplant firms.⁶

The original dataset contains all wholesale trade firms active at some point in the period 1998-2008, yielding 103 251 firm-years for 11 095 firms. The dataset contains firms that are active throughout the period under study, as well as firms entering and exiting during this period. This study uses data on the 4844 firms active throughout the period from the original dataset, while new entrants and exiting firms are excluded. This restriction is imposed by the nature of the evaluation problem – the main question of interest, to investigate profits of relocating firms after relocation, requires that firms can be observed both prior to, and after, relocation.

Table 1 presents descriptive statistics of the data. The first row of the table shows that 2% of all firms were involved in relocation, when relocation is defined as change of the firm's location beyond a municipality border. Column (3) reports mean values for all firms in the sample, columns (4) and (5) report mean values relocating firms and non-

⁶ For multiplant firms, the financial results are aggregated to the headquarters. Thus, while in the data, the location of each single site belonging to a multi-plant firm may be observed, along with some basic characteristics, the annually updated information on firm performance is geographically located to the headquarters of each multi-plant firm.

relocating firms, respectively. These mean values indicate that there are important differences between relocating and non-relocating firms. Relocating firms are, on average, younger and smaller, when compared to non-relocating firms. Furthermore, relocating firms differ in terms of profitability. The average profits, measured as returns on total assets (ROA), are 5.3% for relocating firms, while non-relocating firms have return on assets equal to 8.1%. The differences between relocating and non-relocating firms are in line with the literature on firm relocation determinants (e.g. Hayter, 1997).

Table 1: Variable definition and descriptive statistics

Variable description		Means			
		(3) All firms	(4) Relocating firms	(5) Non-relocating firms	(6) Difference
Relocate _{it}		0.02 (0.0004)			
FirmAge _{it-1}	Number of years since firm's registration	15.50 (0.049)	12.41 (0.31)	15.56 (0.049)	-3.14*** (0.34)
FirmSize _{it-1}	Total sales for firm <i>i</i> in thousands of SEK	25 203.4 (417.4)	19 415.1 (1596.4)	25 325.6 (424.8)	-5910.5** (2933.2)
FirmProfits _{it-1}	Return on total assets for firm <i>i</i>	8.55 (0.08)	5.31 (0.77)	8.62 (0.08)	-3.31*** (0.56)
FirmGrowth _{it-1}	Change in total sales	0.09 (0.001)	0.14 (0.018)	0.08 (0.001)	-0.05*** (0.014)
IndSizeF _{jt-1}	Total sales for industry <i>j</i> in municipality <i>m</i> in thousands of SEK	519 843.8 (3948.8)	619 682.9 (31 270.5)	517 889.5 (3979.1)	101 793.3*** (28 775)
IndSizeS _{jt-1}	Number of firms in industry <i>j</i> and municipality <i>m</i>	27.82 (0.18)	31.95 (1.44)	27.74 (0.18)	-4.20*** (1.36)
MES _{jt-1}	Size of average firm in industry <i>i</i> in municipality <i>m</i>	18 589.9 (132.7)	19 469.2 (1153.8)	18 589.9 (132.7)	901.62 (967.5)
LocComp _{jt-1}	Herfindahl index in industry <i>i</i> and municipality <i>m</i>	0.44 (0.001)	0.42 (0.008)	0.004 (0.001)	-0.02*** (0.008)
PopSize _{mt}	Population size in municipality <i>m</i>	203 137.5 (902.5)	217 144.7 (6913.1)	202 870.3 (910.1)	14 174.4** (6636.2)
PopDens _{mt}	Population density in municipality <i>m</i>	944.92 (4.85)	1172.5 (37.01)	940.5 (4.89)	232.1*** (35.8)
Educ _{mt}	Share of population with university education in municipality <i>m</i>	30.09 (0.33)	32.58 (0.25)	30.04 (0.34)	2.54*** (0.24)
LocGovern _m	Dummy variable, equal to 1 for right-wing controlled municipality board in municipality <i>m</i>	0.27 (0.001)	0.34 (0.11)	0.27 (0.001)	0.06*** (0.01)

Notes: Relocating firms are firms having relocated over a municipality border; Non-relocating firms are firms that have not relocated over municipality borders. Standard errors for the sample means are in parentheses. Statistically significant difference at: *** p<0.01, ** p<0.05, * p<0.1. Industry- and region-specific variables are measured for the out-relocation municipalities.

4.2. Specification of the propensity score model

In the first step of the matching procedure for each firm in the sample its propensity score is estimated. The propensity score corresponds to the firm's probability to relocate given the firm's *ex-ante* characteristics. For this estimation, the following propensity score model is adopted:

$$e(\mathbf{X}, \mathbf{K}, \mathbf{R}) = Pr(\text{Relocation}_{it} = 1 | \mathbf{X}, \mathbf{K}, \mathbf{R}) = \exp(c_0 + \rho' \mathbf{X}_{it-1} + \mathbf{v}' \mathbf{K}_{jt-1} + \gamma' \mathbf{R}_{mt} + \xi_{imt}) \quad (7)$$

The dependent variable in this model, the indicator variable $Relocation_{it}$, identifies for each firm in the sample if the firm belongs to the treatment or control group. The longitudinal nature of the dataset provides a solid background for the identification of relocating patterns of firms, since for each firm-year in the dataset, the location municipality is recorded. Thus, firm i belongs to the treatment group ($Relocation_{it} = 1$), if firm i changes its location address to a different municipality in two consecutive years, otherwise, firm i belongs to the control group ($Relocation_{it} = 0$).

The function of the right-hand side of the propensity score model above is to control for the observable differences among relocating and non-relocating firms. For the quality of the propensity score matching, it is crucial to include in the estimation of the propensity score all variables which might affect participation in the treatment (Heckman et al., 1998). To identify the conditioning variables for the propensity score model, this paper makes use of previous research on firm relocation determinants (see Section 2.2). For simplicity, the conditioning variables are grouped into three vectors: a vector of firm-specific variables \mathbf{X} , a vector of industry location-related variables \mathbf{K} , and a vector of non-industry located related variables \mathbf{R} .

The vector of firm-specific variables, \mathbf{X} , is retrieved directly from the dataset and includes the following attributes for each firm: firm size ($FirmSize_{it-1}$), measured as annual sales in 1000 SEK, firm age ($FirmAge_{it-1}$), measured as number of years since the firm's establishment, firm growth ($FirmGrowth_{it-1}$), measured as the log of the annual sales difference in two consecutive years, and firm profits ($FirmProfits_{it-1}$) measured as the firm's return on total assets.

Besides firm-specific factors, the propensity score model controls for external factors of the geographical environment, where the firm operates. The industry location-related vector \mathbf{K} addresses competitiveness in the sector in which firm i operates, along with the size of the market. The industry-specific variables are compiled by the author to the three-digit level of the NACE classification codes, and geographically to the municipality level. The vector \mathbf{K} includes: market concentration ($MarConc_{jt-1}$), minimum efficient scale (MES_{jt-1}), and two measures of the size of the industry in which firm i operates, industry size ($IndSizeS_{jt-1}$) measured as total sales in industry j and municipality m , and industry size ($IndSizeN_{jt-1}$), measured as the number of firms in industry j and municipality m . The Minimum efficient scale denotes the smallest output a firm should produce to minimize long-term average costs. Following Daunfeldt et al. (2013b), the MES is measured as the size, in terms of sales, of the average firm in industry j and municipality m . The degree of market concentration is measured by a Herfindahl index. The Herfindahl index is the sum of squared market shares of all firms, i , located in municipality m and industry j . The Herfindahl index is defined on the interval 0-1: it has a value equal to 1, if only one firm operates in the local market, and if all firms in the local market have equal size the Herfindahl index is equal to 1/number of firms.

Lastly, the propensity score model controls also for non-industry location-related factors (vector \mathbf{R}), shown in previous research to affect the firm's relocation decision. These variables, retrieved from Statistics Sweden, are measured at the municipality level, and comprise: population size ($PopSize_m$), population density ($PopDens_m$), education level ($Educ_m$), and an indicator variable reflecting type of local government ($LocGovern_m$). The

indicator variable $LocGovern_m$ is equal to one, for right-of-centre local government, and to zero otherwise. Educational level is measured as the share of the population in municipality m with higher (tertiary and above) education. An overview of the variables included in the conditioning vectors, along with variable measurements, is included in Table 1 (first and second column).

In the second step of the matching procedure, relocating firms in the year 2003 are matched with non-relocating firms closest in terms of their propensity scores. The year 2003 is chosen, since it provides an optimal trade-off between the length of the pre-relocation period for the purpose of propensity score matching, and after-relocation period, for evaluation of the effects of the relocation.⁷

The main variable of interest in the second step of the matching procedure is a measure of firm profits. For this purpose, the return on total assets (ROA) is used. ROA measures how much a firm has earned, related to all investments. As emphasized in Libby et al. (2002, p.251), ROA is perceived as the “broadest” measure of profitability and management effectiveness, which is not dependent on the financial strategy of the firm. Simply, firms with higher ROA are doing a better job of selecting and managing investments, *ceteris paribus*. Another option for how to measure firm profits would be to use the firm’s return on equity (ROE). However, compared to ROE, the ROA is a more precise measure than ROE, since ROE is affected by the source of financing.

As mentioned in the previous section, the difference-in-difference estimator (DID) is adopted. The reasoning behind this choice is that even though the author believes that the richness of the dataset addresses the possible issue regarding the conditional ignorability assumption, there might still be a problem of selection on unobservables. A similar strategy might be found in evaluation studies applying a matching framework in the context of human migration (see e.g. Ham et al., 2011; Eliasson et al., 2012). An implication of using the DID estimator is that the dependent variable in the second part of the analysis needs to be defined as prior/after change instead of levels.

Thus, in the main model, the outcome variable of interest in the second part of the analysis is defined as the change in firm profits of firm i :

$$Change\ in\ profit_i = ROA_{i,t+s} - ROA_{i,t-1} \quad (8)$$

where the first term on the right-hand side, $ROA_{i,t+s}$, represents the return on total assets of firm i in year s after relocation, year t , while $ROA_{i,t-1}$ represents the return on total assets of firm i in year $t-1$ prior to relocation. The change in profits is measured in four time periods, the shortest time period is a two-year period, stretching one year prior to the firm’s relocation, and one year after the firm’s relocation ($t-1$ and $t+1$). We also evaluate time periods, starting one year prior to the firm’s relocation and stretching over: two years after the firm’s relocation ($t-1$ and $t+2$), three years after the firm’s relocation ($t-1$ and $t+3$) and four years after the firm’s relocation ($t-1$ and $t+4$), respectively.

⁷ Choosing year 2003, the after-relocation change in profits might be observed in a four-year period. In additional estimations, firms relocating in 2002 and 2004 were considered as treatment groups, the matching procedure yielding qualitatively the same results.

5 Empirical results

The result section of this paper has two parts. In the first part, the results of the propensity score matching are outlined, along with the balancing properties of the matched samples. In the second part, the results of the matching estimators are presented, accompanied by robustness checks.

5.1. Specification and estimates of the propensity score model

The first step in the evaluation procedure is the specification of the propensity score model. Table 2 reports the final specification of the propensity scores model of the firm's decision to relocate. The logit coefficients for the firm-specific variables all have expected signs and most are statistically significant. Consistent with most relocation studies, firm age and firm size are negatively related to the firm relocation propensity, which implies that the firm's probability to relocate decreases with the increasing age of the firm and also that larger firms are less prone to relocate, compared to smaller firms, *ceteris paribus*.

Table 2: Propensity scores estimation

Control variables	Logit	
	Estimates	Standard errors
<i>FirmAge_{it-1}</i>	-0.019***	0.0023
<i>FirmSize_{it-1}</i>	-0.00003***	0.0001
<i>FirmProfits_{it-1}</i>	-0.0002***	0.000
<i>FirmGrowth_{it-1}</i>	-0.01	0.345
<i>MES_{jt-1}</i>	0.000	0.000
<i>IndSizeF_{jt-1}</i>	0.00028	0.0007
<i>IndSizeS_{jt-1}</i>	-0.0002	0.000
<i>LocComp_{jt-1}</i>	0.049	0.104
<i>PopSize_{mt}</i>	-0.01***	0.000
<i>PopDens_{mt}</i>	0.0002***	0.000
<i>Educ_{mt}</i>	0.017***	0.0035
<i>LocGovern_m</i>	0.090	0.061
Constant	-4.563***	0.313
Chi-square statistics	12.42	

Notes: In the estimation of the propensity score model, all firm- and industry-specific variables are lagged by one year. To lag firm- and industry-specific variables is a common way of alleviating a possible reversed causality problem (see e.g. Håkansson et al., 2014; Håkansson et al., 2013; Daunfeldt et al., 2013a). However, in this study, the time lag also reflects the potential migrant firm's decision. The author expects that the firm's decision to relocate is a result of longer considerations, and, thus, is not taken in the same year as the actual firm relocation takes place. The 3-digits NACE coded dummies and year dummies are not reported, but available on request. *** Significant at the 1% level.

The industry-related location variables are not individually significant, but a likelihood ratio test shows that these variables are jointly significant. Lastly, the non-industry location specific variables are all statistically significant, indicating both a positive relationship between population size, population density, and the educational level and the

propensity of the firm to relocate. The results of the propensity score estimations are in line with findings of previous research on firms' relocation determinants in general (Stam, 2007; Knoben et al., 2008; Wetering and Knoben, 2013) and also with studies on the determinants of firms' relocation within the wholesale trade sector (Daunfeldt et al., 2013b; Håkansson et al., 2013).

Table 3 shows the balancing tests for the propensity score estimation, using the relocating firms, and the matched control firms from nearest neighbour matching⁸, using the paired t-statistics for the difference in variables means of relocating firms and the matched sample of non-relocating firms, along with a measure of bias reduction. With the exception of the ($FirmGrowth_{it-1}$) variable, the test statistics are not near statistical significance at the standard test level of 5% or 10% for any of the variables controlled for in the propensity score matching method, indicating that the matching procedure balanced the pre-relocation characteristics between relocating firms, and the matched group of non-relocating firms.

Table 3: Balancing tests

Variable	Mean values		% bias	% bias reduction	Paired <i>t</i> statistics
	Relocating firms	Non-relocating firms			
$FirmAge_{it-1}$	15.11	13.79	10.1	42.7	0.72
$FirmSize_{it-1}$	21 540	24 138	-3.3	70.6	-0.36
$FirmProfits_{it-1}$	82.4	97.1	-8.8	-87.2	-0.54
$FirmGrowth_{it-1}$	0.17	0.02	23.7	-17.2	1.72
MES_{jt-1}	15 550	13 301	8.5	30.5	1.08
$IndSizeF_{jt-1}$	43.4	48.0	-7.7	56.4	-0.44
$IndSizeS_{jt-1}$	720 000	670 000	4.8	71.8	0.28
$LocComp_{jt-1}$	0.37	0.32	15.5	-35.7	0.95
$PopSize_{mt}$	280 000	250 000	13.1	51.8	0.72
$PopDens_{mt}$	1480.2	1292.2	12.5	69.1	0.09
$Educ_{mt}$	33.1	33.2	1.4	96.3	0.08
$LocGovern_m$	0.27	0.35	-19.4	-54.8	-1.11

5.2. Matching estimator results

The second step of the evaluation procedure consists of matching the treatment and control units and comparing the mean outcome for both groups. In the matching procedure, a nearest-neighbour matching method is used. In nearest neighbour matching, a fixed number of neighbours is defined, and observations with the closest propensity scores form the matched control group.

Table 4 presents matching results, using the previously estimated propensity scores and considering firm relocation to be equal to a treatment. As a treatment group, firms having changed the location of its business beyond a municipality border in the period 2003 are considered, while the matched control group is constructed from non-relocating firms in

⁸ The results are similar for nearest neighbour matching with two, three, and four nearest neighbours.

the period 2003, without previous relocation history.⁹ The outcome variable of interest in the matching estimator is the change in profits, where profits are measured as the firm's return on total assets. The difference in mean outcomes between relocating firms and non-relocating firms is evaluated over four different time periods, as mentioned above. Columns (2) and (3) in Table 4 report the mean outcome for the treated and control groups, while the average treatment effects, standard errors, and t-statistics are reported in columns (4), (5) and (6), respectively.¹⁰

As pointed out in di Cintio and Grassi (2010), when interpreting the results of the matching estimator, it is important to keep in mind that in this case the control group was constructed to simulate the missing potential outcome of relocating firms, and, thus, the estimates pertain only to the gains/losses experienced by relocating firms. As such, the estimates of the average treatment effect on the treated (ATT) in Table 4 cannot be interpreted as pure differences among relocating and non-relocating firms; rather the ATT should be interpreted as the change in profits caused by the firm's relocation.

Panel A of Table 4 summarizes the results, using nearest neighbour matching with one neighbour. In this matching procedure, known also as one-to-one matching, the matched control group is constructed from a set of non-relocating firms, using only one non-relocating firm, closest in the probability to relocate, for each relocating firm. The results indicate that in the shortest evaluated period spanning one year prior to, and one year after, the firm's relocation, relocating firms experience an increase in mean profits with 2.14 percentage points. Meanwhile, the matched control group, corresponding to the missing potential outcomes of relocating firms, experienced in the same time period, decrease in mean profits with -6.23 percentage points, resulting in a positive and statistically significant average treatment effect (ATT) of the relocation. From an economic point of view, this result means that relocating firms experience profits gain after relocation, compared to what the profits would be, had they not relocated, and this increase is on average 8.37 percentage points, compared to the pre-relocation profits of relocating firms. A positive ATT is observable also once the profit change between relocating and non-relocating firms is evaluated for a time period spanning over two, three, and four years after the firm's relocation. The ATT varies between 5.27 percentage points in the two years after relocation, and 9.98 percentage points, four years after relocation. However, judging from the t-statistics, the ATT is statistically significant only for two out of the four analysed time spans.

Panel B presents the results of the nearest neighbours matching, using two nearest neighbours. Panel C presents the results using three nearest neighbours, while Panel D contains the results utilizing with four nearest neighbours. In other words, this means that

9 As mentioned previously, the data covers the period 1998-2008. The year 2003 is chosen for evaluation of the effects of relocation, since this provides an optimal trade-off between the length of the pre-relocation period for the purpose of matching, and after-relocation period for evaluating the effects of relocation. However, in additional estimations, the relocating firms in years 2002 and 2004 were examined, yielding qualitatively similar results.

10 The estimation procedure is carried out by using the `psmatch2` Stata command, developed by Leuven and Sianesi (2003). The `psmatch2` routine calculates the standard errors of the treatment effects, assuming independent observations, homoskedascity of the outcome variable within the treated and control groups, and that the variance of the outcome not to depend on the propensity scores. Based on Imbens (2004), bootstrapping is not used to calculate the standard errors.

the matched control group of firms is constructed from a set of all non-relocating firms, using the two, three, respective four firms closest in proximity to relocate, for each relocating firm in the sample.

Table 4: The matching results

Treatment	Firm relocation beyond municipality border in 2003				
Outcome	Firm profits change (percentage points) for one, two, three and four years after relocation				
Treatment group	Relocating firms in 2003				
Control group	Non-relocating firms in 2003				
Matching algorithm	Nearest neighbour matching				
	(2)	(3)	(4)	(5)	(6)
	Treated	Control	ATT	S.E.	t-value
<i>Panel A: Nearest neighbour (one-to-one matching)</i>					
Change in profits $t-1$ and $t+1$	2.14	-6.23	8.37***	3.31	2.53
Change in profits $t-1$ and $t+2$	2.58	-2.68	5.27	3.70	1.42
Change in profits $t-1$ and $t+3$	4.90	-5.07	9.98**	4.90	2.04
Change in profits $t-1$ and $t+4$	5.32	-5.87	11.19***	1.43	2.52
<i>Panel B: Two nearest neighbours</i>					
Change in profits $t-1$ and $t+1$	2.14	-5.08	7.22***	3.01	2.41
Change in profits $t-1$ and $t+2$	2.58	-2.66	5.25	3.21	1.64
Change in profits $t-1$ and $t+3$	4.90	-4.38	9.29**	3.92	2.37
Change in profits $t-1$ and $t+4$	5.32	-5.05	10.38***	3.70	2.80
<i>Panel C: Three nearest neighbours</i>					
Change in profits $t-1$ and $t+1$	2.14	-2.45	4.52	2.91	1.58
Change in profits $t-1$ and $t+2$	2.58	-6.50	3.23	3.08	1.05
Change in profits $t-1$ and $t+3$	4.90	-2.77	7.68**	3.73	2.06
Change in profits $t-1$ and $t+4$	5.32	-2.15	7.47**	3.40	2.15
<i>Panel D: Four nearest neighbours</i>					
Change in profits $t-1$ and $t+1$	2.14	-2.50	4.64	2.83	1.64
Change in profits $t-1$ and $t+2$	2.58	-0.68	3.25	2.93	1.11
Change in profits $t-1$ and $t+3$	4.90	-2.04	6.95*	3.56	1.95
Change in profits $t-1$ and $t+4$	5.32	-2.30	7.69**	3.35	2.27

Note: The estimation procedure is performed with a user-written STATA command *psmatch2* developed by Leuven and Sianesi (2003). Common support condition is imposed by “*common*” option in the specification of the matching command. To reinforce the common support constraint, in line with Heckman et al. (1998), trimming level=3 is used. Standard errors on the treatment effects assume independent observations, homoskedascity of the outcome variable within the treated and control groups, and the variance of the outcome not to depend on the propensity scores. The estimated propensity score is based on the logit estimation of Eq. (7).

Thus, the size of the matched control group in these additional specifications is larger than the size of the matched control group in Panel A. The main pattern in these three additional matching specifications is similar, as in Panel A, using one-to-one matching – the ATT of relocation is positive for all evaluated time periods, indicating that relocating

firms experience profit gain, compared to what the profits would be, had the firms not relocated. The size of the ATT varies between 3.2-10.3 percentage points, and in seven out of twelve estimations, the resulting ATT are statistically significant at the conventional level of significance.

As noted in Section 3, the second assumption of the matching estimator requires matching to be used only over the portion of data where each treatment unit can find a reasonable number of control units. To ensure that firms similar in terms of their covariates are compared, in Table 4, the common support condition was imposed by dropping treatment observations, whose estimated propensity score is higher than maximum, or less than minimum of the propensity scores of the control observations. In addition to that, the procedure proposed originally in Heckman (1998) is followed, and 3 % (trimming level $q=3$) of relocating firms, for which the propensity score density of the control observations is the lowest, were dropped. To check the sensitivity of the matching estimator to the trimming level, a trimming level of $q=5$ and $q=7$ was considered in additional estimations. Replicating Table 4, using a trimming level of 5 % or 7 % does not alter the average treatment results. Similarly to Table 4, all estimated ATT are positive, and eight out of twelve are statistically significant at the conventional level of significance.

Another feature of nearest-neighbour matching (see Stuart, 2010) is the potential problem of poor matches, a situation which occurs when the nearest neighbours are far removed from the treatment units, in terms of estimated propensity scores. To avoid this problem in the matching procedure, a tolerance level on the maximum of propensity score distance using a caliper was imposed. The caliper restricts the matching procedure to those matches, having sufficiently small difference in propensity scores between the treated and untreated firms. The main model presented in Table 4 relies on a caliper of 0.05, i.e. only control units having a maximum of a 5 % difference in the propensity score from the treated firms are allowed as controls. In additional estimations, testing was done whether using a different caliper changes the results. Replicating the matching procedure in Table 4, using calipers of 1 %, or 2.5 % do not, in any major way, alter the average treatment effects of relocation, or the statistical significance of the ATT estimates.

6 Conclusions

In recent decades, firm relocation has received an increasing interest in research literature, and firm relocation determinants mainly have come under scrutiny. In studies within the neoclassical framework, a firm's relocation decision is frequently viewed as a decision based on whether there are expected net gains in profits as a result of the relocation. However, only a few authors examine empirically if this basic assumption on profit gains of relocation holds. With this background, the purpose of this article is to empirically test if relocating firms increase their profits in the period after relocation, compared to what the profits would be, had not the firms relocated. To this end, a longitudinal dataset of Swedish limited liability firms in the wholesale trade sector, and difference-in-difference propensity score matching is used. For each firm in the dataset, a propensity score corresponding to the probability to relocate given the baseline characteristics of the firm

was calculated, and treatment (relocating) firms were paired with control group (non-relocating) firms, with similar propensity scores. Using propensity scores, the presumed difference between relocating and non-relocating firms, reported frequently in firm relocation research, was controlled for and, thus, similar groups of firms were compared. In order to also address possible selection on unobservables, an issue often reported in studies of human migration and post-migration earnings, the propensity score matching was combined with a difference-in-difference estimator. The matching procedure was conducted, using nearest-neighbour matching. In the main model, each relocating firm was compared with the one non-relocating firm closest, in terms of the estimated propensity score, i.e. having the most similar baseline characteristics in the pre-relocation period. In additional matching procedures, the two, three, and four non-relocating firms with the closest propensity scores were also used to construct the matched control group. The results of the main model show that relocating firms experience an average increase in profits in the range of 5 to 11 percentage points, due to relocation, depending on the length of the time span after the relocation in which the ATT is measured. The resulting ATT is statistically significant at conventional levels in the majority of estimations.

The presented results broadly support the findings in Knoben et al. (2008), where the effects of firm relocation on firms in the Dutch automation industry were examined. The authors report that relocating firms experience a decrease in profits in two years after firm relocation, while in a five-year period after relocation there is a positive effect of relocation upon the firm's profitability. Compared to the study of Knoben et al. (2008), the longitudinal nature of the dataset utilized in this study makes it possible to evaluate the impacts of relocation in a continuous time period after relocation, rather than for selected years only. Furthermore, the firm-level data set covering all limited liability firms makes it possible to address a potential sample selection bias due to missing data for several firms, a possible limitation of the study performed by Knoben et al. (2008), where only a minority of the sampled firms answered the questionnaire sent out by the researchers. Lastly, the detailed information on each firm, originating from the firms' financial sheets makes it possible to address self-selection into relocation. Self-selection, in the context of relocation, implies that certain type of firms, regardless whether they are 'well-performing', or firms with low profitability, move more frequently than other types of firms do. If the self-selection factor is related not only to the relocation decision, but also to the outcome of relocation, in this case the firm profits, then not addressing this self-selection may yield biased results of the estimations. However, as mentioned earlier, the estimation strategy in the form of propensity score matching, commonly applied within human migration research (e.g. Eliasson et al., 2012) and occasionally also using firm-level data (e.g. Wagner, 2010), is one way of controlling for the potential bias originating from self-selection on observable factors. Combining this with a difference-in-difference estimator also holding constant for all time-invariant heterogeneity between relocating and non-relocating firms makes a strong case that the findings reported in this study are the effects of the firm's migration decision on the profits of the relocating firm.

Although the findings of this paper provide new and valuable insights, the study has some limitations. Firstly, the focus of the study on a single sector in a single country limits the generalizability of the results. The main motivation for the focus in this paper on

wholesale trade was a high relocation frequency for this sector, reported in studies conducted in Sweden and other European countries, giving a sample of relocating firms large enough to make statistical inference. The main requirements of the firms in this industrial sector are a good infrastructural accessibility to the market, and suitability of sites to store and distribute large amount of goods on a daily basis. On the other hand, this industrial branch is not characterized by the high sunk-costs as many of the branches in the manufacturing sector are, and, therefore, wholesaling firms are expected to be more 'mobile' compared to, for example, the majority of branches in the Swedish manufacturing sector. Even though this is the case, the relocation frequency in the sample is on average only 2%, i.e. only two out of hundred wholesale firms change the location of their business annually. Given that, the author considers that choosing other industrial sectors with an even lower relocation frequency would mean drawing inference from so few relocating firms that the results could not be trusted to reflect the actual effects of relocation on firm profits. Thus, this article opts for conducting a single sector analysis only.

Secondly, the estimation strategy in the form of propensity score matching, combined with a difference-in-difference approach, disqualified from the estimations all non-surviving firms, since for each firm, observations of the firm's characteristics both in the period prior to and after relocation are required. This implies that firms that relocated, but did not survive the relocation and closed down shortly after relocation, are not included in the estimations. This may lead to the conclusion that relocation has more positive effects on profits of firms than is the case, and the results should thus be interpreted as an upper bound of the positive effects of relocation on firm profits. Moreover, the analysis is focused solely upon single establishment firms, leaving the 14% of firms in the dataset with more than one establishment out of the estimations. Even though the findings of previous studies show that multi-establishment firms usually adopt different strategies in situations when single establishment firms would instead relocate (Pellenbarg et al., 2002), and the share of the multi-plant establishment firms among the wholesale trade firms in the dataset is rather low, inference should still only be drawn for single-plants establishments.

Literature

- Acs, J.Z., D.B. Audretsch, P. Braunerhjelm and B. Carlsson. 2009. The knowledge spillover theory of entrepreneurship. *Small Business Economics*. 32: 15-30.
- Arauzo-Carod, J.M. 2005. Determinants of industrial location: An application for Catalan municipalities. *Papers in Regional Science*. 84: 105-120.
- Baumol, W.J.1990. Entrepreneurship: Productive, Unproductive and Destructive. *Journal of Political Economy* 98: 893-921.
- Becker, S.O. and A. Ichino. 2002. Estimation of average treatment effect based on propensity scores. *Stata Journal*. 2(4): 358-377.
- Bernini, C., and G. Pellegrini. 2011. How are growth and productivity in private firms affected by public subsidy? Evidence from regional policy. *Regional Science and Urban Economics*. 41: 253-265.
- Bodenman, B.R. and K.W.Axhausen, 2010. Destination choice for relocating firms: A discrete choice model for the St.Gallen region, Switzerland. *Papers in Regional Science*. 91(2) 321-341.
- de Bok D. and van F. Ort. 2011. Agglomeration economies, accessibility, and the spatial choice behaviour of relocating firms. *The Journal of Transport and Land Use*. 4(1): 5-24.
- Brouwer, A.E., I.Mariotti and J.N. Ommeren. 2004. The firm relocation decision: An empirical investigation. *The Annals of Regional Science*. 38(2): 335-347.
- Bryson, A., R. Dorsett and S.Purdon. 2002. The use of propensity score matching in the evaluation of active labour market policies. Research working paper, 4. Department for Work and Pensions, London, UK.
- di Cintio, M, and E.Grassi. 2013. Internal migration and wages of Italian university graduates. *Papers in Regional Science*. 92: 120:140.
- Daunfeldt, S-O., N. Elert and N. Rudholm. 2013a. New start-ups and firm in-migration: evidence from the Swedish wholesale trade industry. *Annals of Regional Science*. 51: 479: 494.
- Daunfeldt, S-O., Å. Lang, Z. Macuchova and N.Rudholm. 2013b. Firm growth in the Swedish retail and wholesale trade industries. *Service industries journal* 33: 1193-1205.
- van Dijk, J. and P.H. Pellenbarg, 2000. Firm relocation decision in Netherlands: An ordered logit approach. *Papers in Regional Science*. 79: 191-219.
- Eliasson, K., R. Nakosteen, O.Westerlund and Zimmer M. 2012. All in the family: Self-selection and migration by couples. *Papers in Regional Science*. 93(1): 102-124.
- Figueiredo, O., Guimarães, P. and D. Woodward. 2002. Home-field advantages: Location decision of Portuguese entrepreneurs. *Journal of Urban Economics*. 52: 341-361
- Garnsey, E. 1998. A Theory of the early growth of the Firm. *Industrial and corporate change*. 7: 523-555.
- Imbens, G.W. 2004. Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*. 86: 4-29.
- Kauder, B. 2014. Incorporation of municipalities and population growth: A propensity score matching approach. *Papers of Regional Science*. doi: 10.1111/pirs.12148
- Knoben, J. and Oerlemans, L.A.G., 2008. Ties that Spatially Bind? A Relocational Account of the Causes of Spatial Firm Mobility. *Economic geography*. 84(2): 158-182.

- Knoben, J., Oerlemans, L.A.G. and Rutten R.P.J. H., 2008. The effects of spatial mobility on performance of firms. *Economic geography*. 84(2): 158-182.
- Knoben, J. 2011. The geographical distance of relocation search: An extended resource-based perspective. *Economic geography*. 87(4): 371-392.
- Kronenberg, K. 2013. Firm relocation in Netherlands: Why do firms move, and where do they go? *Papers in Regional Science*. 92(4): 692-713.
- Libby, R., R. Bloomfield and M.W. Nelson. 2002. Experimental research in financial accounting. *Accounting, organization and society*. 27: 775-810.
- Ham, J.C., L.Xianghong and P.B. Reagan. 2011. Matching ad semi-parametric IV estimation, a distance-based measure of migration and the wages of young men. *Journal of Econometrics* 161: 208-227.
- Hayter, R. 1997. *The Dynamics of Industrial Location. The Factory, the Firm and the Production System*. New York: Willey
- Heckman, J.J., Ichimura, H. and Todd, P. 1998. Matching as an econometric evaluation estimator. *Review of Economic Studies*. 65: 261-294.
- Holl, A. 2004. Start-ups and relocations: Manufacturing plant location in Portugal. *Papers in Regional Science*. 83 (4): 649-688.
- Håkansson, J., Z. Macuchova and N. Rudholm (2013). Firm migration in the Swedish wholesale trade sector. *International Review of Retail, Distribution and Consumer Services*. 23: 48-64.
- Marshall, A. 1920. *Principles of Economics*. Macmillan, London.
- Nakosteen, R. A. and M.A. Zimmer. 1987. Determinants of regional migration of manufacturing firms. *Economic inquiry*. 15: 351-362.
- Pellenbarg, P.H., L.J.G. van Wissen and J. van Dijk. 2002. Firm Migration. In *Industrial Location Economics*, ed. P.McCann, 110-148. Cheltenham, UK: Edward Elgar Publishing
- Stam, E. 2007. Why butterflies don't leave: Locational behaviour of entrepreneurial firms. *Economic geography* . 83: 27-50.
- Stuart, E. A., 2010. Matching methods for causal inference: A Review and a Look forward. *Statistical Science*. 25 (1): 1:21.
- Rubin, D.B. and N. Thomas. 1996. Matching using estimated propensity scores, relating theory and practice. *Biometrics*. 52: 249-264.
- Rosenbaum, P.R., and D.B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 70: 41-55.
- Sleutjes, B. and Völker, B. 2012. The role of the neighbourhood for firm relocation. *Tijdschrift voor Economische en Sociale Geografie*. 103: 240-249.
- Wagner, J. 2011. Offshoring and firm performance: self-selection, effects on performance or both? *Review of World Economics*. 147: 217-247.
- Wetering, A. and Knoben, J. 2013. Footloose: An analysis of the drivers of firm relocations over different distances. *Papers of Regional Science*. 92(4): 792-809.
- Wenz, M. 2008. Matching estimation, casino gambling and the quality of life. *Annals of regional science*. 42: 236-249.