RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

IFN Working Paper No. 1522, 2025

The Role of Specialized Knowledge and 'Know-How' for Firm Productivity: Evidence from the Equine Industry

Pia Nilsson, Hans Andersson and Tobias Heldt

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

The role of specialized knowledge and 'know-how' for firm productivity: evidence from the equine industry

Pia Nilsson[1],[2], Hans Andersson[2], Tobias Heldt[3]
[1]Research Institute of Industrial Economics (IFN)
[2]Swedish University of Agricultural Sciences
[3]Center for Tourism and Leisure Research, Dalarna University[1]

Abstract

Using matched employer–employee data, we investigate the influence of human capital inputs on firm productivity. Several variables are used to measure firms' access to skilled labor, such as their share of employees with occupation-specific education and experience in horse breeding and hippology and access to a local pool of skilled labour. The results show that occupation-specific training is associated with an average productivity premium of 11%, but there is significant intra-industry heterogeneity in the extent that firms that can gain from workers with specialized training. The results have implications for policy and the investment decisions made by firms.

Keywords: Human capital, Skills, Productivity, Horse industry
JEL classification: J24, D22, D24

[1]Funding from The Swedish-Norwegian Foundation for Equine Research project H-20-47-566 is gratefully acknowledged.

1 Introduction

The influence of human capital inputs, such as experience and education, on firm productivity and growth has received a lot of attention in the literature (Black & Lynch 1996, Glaeser & Maré 2001, Fox & Smeets 2011, Almeida & Carneiro 2009, Konings & Vanormelingen 2015, De la Roca & Puga 2017, Serafinelli 2019, Crescenzi & Gagliardi 2018, Morris et al. 2020). Within this large literature, there is evidence that the characteristics of the workforce, such as their type and level of education and labour market experience, play a significant role in explaining firm performance. Despite this progress, important gaps remain to be explored. Most of the previous studies have focused on firms in non-agrarian industries, such as manufacturing, retail, and knowledge-intensive business sectors (e.g. Parrotta et al. 2014, Barzotto & De Propris 2019, Audretsch & Belitski 2023), and evidence of rural industries is rare. Existing studies with an agricultural focus have also mainly examined only one dimension of human capital approximated by the level and type of education of managers (Sumner & Leiby 1987, Asadullah & Rahman 2009, Reimers & Klasen 2013, Nowak & Kijek 2016), leaving the productivity gains associated with other human capital inputs and the workforce largely undetermined.

In this paper, we provide novel evidence from a rural industry where specialized skills and tacit knowledge are central components of the production process. A particularly interesting feature of the equine industry is its labor intensity and reliance on specific skills and know-how (Jez et al. 2013). Horse industry related demand and supply also have a tendency to cluster spatially, and the products and experiences provided by firms are often dependent on a complex set of specialized suppliers (McManus et al. 2013). This suggests that there should be potential for productivity gains arising from occupation-specific training and experience, as well as the movement of knowledgeable workers across firms. However, we are unaware of any studies that have addressed the potential productivity gains arising from human capital inputs in the industry, making their nature and significance largely unknown. The purpose of this paper is to investigate whether there exist productivity gains derived from increases in the employment share of workers with formal training of direct relevance for an occupation in the industry. We also examine whether related educational qualifications and educational diversity matter for firm productivity and, if so, whether firms can gain from having workers with similar or complementary skills. We conduct the first population study on the productivity effects associated with skilled workers in the industry by linking registry data for all active firms and their employees over a 13-year period. The data allow us to observe details on the educational qualifications of workers in related and unrelated fields, making our industry approach suitable for testing hypotheses on the role of general and specific human capital in firm productivity (Becker 1964, Kremer 1993, Noteboom 2000, Lazear 2009). Our intra-industry approach allows us to overcome biases related to technology choice and the common finding that skill-weights can vary significantly even within narrowly defined industries (Rigby & Essletzbichler 2006, Lazear 2009). Another advantage is that we can observe workers' employment histories and precisely measure the accumulated stock of occupation-specific experience of all employees and managers belonging to a firm. We can thus reduce biases related to tenure and contribute to the literature that focuses on the productivity gains associated with the firm's stock of experience (Crescenzi & Gagliardi 2018, Audretsch & Belitski 2023). Firm-level data comprise information on the gross value of production, inputs employed in production, and location of the firm, which means that we can augment the production function to account for local characteristics of the labor market, such as the access of firms to specialized suppliers and a local pool of workers with the relevant skills (Marshall-Arrow-Romer (MAR)-type externalities). We can thereby account for that firms with more internal knowledge and "absorptive capacity" are better at exploiting external knowledge, and that the two sources of knowledge can complement each other (Cohen & Levinthal 1990).

Our production function estimation is based on control function approaches that account for several potential sources of endogeneity, such as simultaneity in inputs choice, panel attrition resulting from exits and labour market frictions (Levinsohn & Petrin 2003, Ackerberg et al. 2007, 2015, Gandhi et al. 2020, Hu et al. 2020). The analysis relates implied total factor productivity to different types of labour as inputs and allows for worker heterogeneity along additional dimensions, such as age, gender and education level. The results are consistent with theories of human capital and productivity with the implication that profit-maximizing firms in the industry should increase workers with occupation-specific training and experience. Examinations of intra-industry heterogeneity, however, reveals large differences across firms in the industry, and it is only firms specializing in horse racing that can benefit from increases in workers with occupation-specific training in horse breeding and hippology. The results are robust to IV estimations, where we instrument specialized training with indices that measure the predicted stock of workers with such training at the level of the local labour market (LA) based on the stock observed in 1993. We also find some evidence pointing towards potential productivity gains of co-location, suggesting that there exist strategic complementarities between firms in the industry. Although comparable approaches have been used to determine the productivity gains associated with skilled labour in non-agrarian industries (e.g. Parrotta et al. 2014, Konings & Vanormelingen 2015, Boikos et al. 2023, Braunerhjelm & Lappi 2023), we are unaware of studies that have tested such hypotheses in an agrarian context focusing on firms in the horse industry.

2 Background

While the overall performance of the agricultural sector in Europe has seen a declining trend in the past decades, the equine industry is growing. Recent statistics show that there are more than 6 million horses that graze on 6 million acres of permanent grasslands, and the number of riders increases by around 5% each year.[2] Structural transformation in agriculture in combination with increased leisure is expected to reinforce these trends, which has led to a greater focus on the potential of the industry to contribute to rural and regional growth and preservation of permanent grasslands (Zasada et al. 2011, López et al. 2019). These trends are not distinctive features for the Swedish and European context, but are found in countries such as Australia, New Zeeland, the UK and the US (Pickering et al. 2009). The structure of the horse industry in an international context has been thoroughly explored by McManus et al. (2013), who especially points out the link between breeding and racing, and the complex relationships that exist between firms in the industry. The logic for the industry of horse racing is similar to professional sports, where expansion is driven by "new money" coming in via the purchase of a horse put into professional training, ie firm output is price money from racing and training fees. An increase in demand for horse racing, both the total number of horses bought and sold and the value of horses at the top-end of the market, has been described as resulting in a trickle-down of price increases through lower levels of the market. A large part of the business model to attract "new money" is the narrative that anyone can be the owner of a successful racehorse, but there is a higher probability that a higher-priced horse earns higher prize money over a racing career than a cheaper purchase. Input goods and services for success in horse racing are, apart from the horse itself, quality in factors feed, veterinary attention, training, stabling and other aside services like equine therapy and finally factors related to the trainer itself. This includes knowledge acquired through experience and activities that are needed to perform well in racing. Especially the training and caring (including feed) for the horse is both labor and knowledge intensive.

The logic of horse breeding is more closely tied to agriculture as it requires similar types of input and facilities and builds on the knowledge and experience used to feed and raise other animals. The output of horse breeding is naturally a foal, which is usually sold at auctions, either yearling sales or after some extra months of training, but the breeder can choose if and when to sell depending on preferences and the current market price. Despite that both horse breeding and horse racing are becoming increasingly important parts of rural and regional economies (Heldt et al. 2018), there is little empirical evidence of the drivers of productivity in the industry.

In low-tech industries dominated by small firms, the learning that is important to productivity

[2] The European Horse Network: Horses for Growth and Environment (2015)

is often the type that takes place when workers handle and develop day-to-day routines (Malerba 1992). In labor-intensive sectors, productivity is closely linked to the skills embodied in workers and firm managers (Boschma et al. 2009). However, the type of qualification that matters is not clear, as this can vary across industries. An important source of knowledge is acquired through work experience that can increase the productivity of individuals as they can learn new skills and perfect previous skills through on-the-job training (Becker 1964). Another important source of knowledge is related to the level of education and the area in which workers have been formally trained. A question that has raised debate in the literature is whether firms should combine workers with specialized skills or invest in a workforce with complementary skills to promote productivity and technical efficiency (Iranzo et al. 2008, Orr 2022, Cunningham et al. 2023).

Economic theory suggest that firms can gain productivity from having workers with occupation-specific skills, but combinations of workers with varying skills and educational qualifications have also been shown to facilitate firm productivity. Diversity in skills can provide informational and organizational advantages and improve firms' capacity to adopt new technologies (Penrose 2009, Østergaard, Timmermans & Kristinsson 2011). Educational diversity can also generate knowledge spillover effects within and across firms provided that workers' knowledge sets are not too cognitively far apart (Noteboom 2000). An additional common finding in the literature is that firms with more internal knowledge and 'absorptive capacity' are better at exploiting and taking advantage of external knowledge, suggesting that there may exist complementarities between firms' access to internal and external knowledge (Cohen and Levinthal 1990; Backman, 2014).

The implications of workers skill sets for firm productivity and growth have received considerable attention in the literature and the empirical evidence is mixed. The paper by Iranzo et al. (2008) study Italian firms in manufacturing and show a positive relationship between firm productivity and skill dispersion. They find this to hold within, but not between firms. The study by Östbring and Lindgren (2013) on Swedish employer–employee data shows that a high degree of related knowledge in the workforce has a strong positive effect on firm performance and mainly in labour-intensive industries. Wixe and Andersson (2017) use comparable Swedish data and show that cognitive relatedness in terms of sharing a common educational background is positively associated with firm productivity.[3] As stated in the introduction, and as a motivation for this paper, these hypotheses have not been examined in empirical analyses of firms productivity in the equine industry.

[3] Parrotta et al.(2014) use employer-employee matched data on Danish firms in manufacturing, retail, and business services and show mixed evidence regarding the role of diversity in education on firm productivity across industries.

3 Methods and data

3.1 Estimation strategy

Based on our research purpose, we use a gross-value based estimate of Total Factor Productivity (TFP) to examine the firm-level relationship between productivity and skilled labour. A gross-output based estimate of TFP includes intermediate inputs and thereby provides a more complete picture of the production process compared to a value-added based estimate (Christensen 1975). This output measure is well suited for our industry approach as firms in the equine industry rely heavily on intermediate inputs, such as feed and fodder (Heldt et al. 2018). TFP measures the efficiency of which firms convert inputs into outputs and has been widely applied to study the relationship between human capital inputs and firm productivity (Parrotta et al. 2014, Konings & Vanormelingen 2015, Braunerhjelm & Lappi 2023). However, consistent estimation of TFP is hampered by several methodological challenges, such as simultaneity bias, firm exits, and adjustment frictions in labour inputs (Marschak & Andrews 1944, Olley & Pakes 1996, Levinsohn & Petrin 2003, Ackerberg et al. 2007, 2015). To address these potential sources of endogeneity, we specify the following relationship between output, skilled labour and the efficiency level of firms in a Cobb-Douglas framework:

$$y_{ijt} = \alpha + \gamma_l l_{ijt} + \gamma_k k_{ijt} + \gamma_m m_{ijt} + \omega_{ijt} + \eta_{ijt} \tag{1}$$

where y_{ijt} measures gross-output of firm i in sub-industry j, which is either horse racing or horse breeding, at time t and where l_{ijt} is a vector of variables that measure skilled and unskilled labour inputs constructed from headcounts in our employer-employee matched data. Capital and intermediate inputs are measured by k_{ijt} and m_{ijt} and the production function relates output to inputs and the efficiency level of firms A, such that $\ln A_{ijt} = \alpha + \epsilon_{ijt}$ where $\epsilon_{ijt} = \omega_{ijt} + \eta_{ijt}$.[4] The model has two unobservable terms, the first is a residual η_{ijt} and the second is the productivity of the firm ω_{ijt}, which is assumed to follow the following first-order Markov process:

$$\omega_{ijt} = \mathbb{E}(\omega_{ijt} \mid \omega_{i,j,t-1}) + \xi_{ijt} = g(\omega_{i,j,t-1}) + \xi_{ijt} = g(\phi_{t-1} - \alpha - \gamma_k k_{i,j,t-1}) + \xi_{ijt} \tag{2}$$

where ξ_{ijt} represents an innovation term (Olley and Pakes, 1996). The difference between the two is that while the former is assumed to be uncorrelated with firms' period t input choices, the latter can affects such choices and the fact that ω_{ijt} is unobservable and potentially influential can lead to simultaneity bias in production function estimations (Marshak and Andrews, 1944). To address this, we build on the method proposed by Levinsohn and Petrin

[4] We follow the denotations used in this literature (e.g. Ackerberg et al., 2015) and use lowercase letters denote the log of a variable.

(2003) (henceforth LP) in using firms demand for intermediate inputs to define a control function for ω_{ijt}.[5] We assume that inputs are either variable (intermediates) or quasi-fixed

[5]The alternative approach to use firms demand for investments as in Olley and Pakes (1996) is not an option for us as firms in our data frequently report zero investment, which would exclude a large number of firms in the analysis. Nearly all firms in our data report positive values on intermediate inputs and we therefore build on the approach by Levinsohn and Petrin (2003).

(capital) and that capital accumulation follows a law of motion such that firms capital stock in t is determined by the investments made in *t-1*, assuming that it is uncorrelated with η_{ijt}. Profit-maximizing behavior will thus lead firms to increase their use of variable inputs in response to a positive productivity shock occurring in t, but not of quasi-fixed inputs because these decisions were made in the previous period. We can thereby express firms demand for intermediate inputs as a function of the state variables k_{ijt} and ω_{ijt}. In addition, and given that demand for intermediate inputs is strictly monotonic in ω_{ijt} and has only one unobservable among its arguments (scalar unobservable assumption), we can invert the demand function to obtain the following control function:

$$\omega_{ijt} = h_t(k_{ijt}, m_{ijt}) \tag{3}$$

Substituting equation 2 into the production function (equation 1) provides the first-stage equation in our control function estimation:

$$y_{ijt} = \gamma_l l_{ijt} + \phi_t(k_{ijt}, m_{ijt}) + \epsilon_{ijt} \tag{4}$$

where $\phi_t(\cdot) = \alpha + \gamma_k k_{ijt} + \gamma_m m_{ijt} + \omega_{ijt}(k_{ijt}, m_{ijt})$. In our empirical application, we estimate the first-stage equation using OLS and a third-order polynomial approximation using k_{ijt} and m_{ijt} in place of $\varphi(\cdot)$. Furthermore, and since our production function is gross-output with intermediate inputs as proxy, we use a generalized method of moments estimator in a second stage to identify the input coefficients and control for the endogeneity of the variable inputs. We specify the following moment conditions

$$\mathbb{E}[\epsilon_{ijt} + \xi_{ijt} \mid k_{ijt}, m_{i,j,t-1}] = 0 \tag{5}$$

and calculate firm total factor productivity as a residual using the production function estimates:

$$\text{TFP}_{ijt} = \widehat{\epsilon_t + \xi_t} = y_{ijt} - \hat{\gamma}_l l_{ijt} - \gamma'_k k_{ijt} - \gamma'_m m_{ijt} - \mathbb{E}(\widehat{\omega_t \mid \omega_{t-1}}) \tag{6}$$

where $\hat{\omega}_{ijt} = \hat{\phi}_{ijt} - \gamma'_k k_{ijt} - \gamma'_m m_{ijt}$.

A potential limitation of this approach for our purpose is that the labor coefficient may not be identified in the first-stage estimation if there exists adjustment frictions, such as hiring, firing and search costs, that prevent firms from changing labor inputs in response to changes in ξ_{ijt} (Bond & Söderbom 2005, Ackerberg et al. 2007, 2015). Such adjustment frictions are likely present in a context such as Sweden which is governed by collective agreements and strict labor market legislation. They are also likely to be particularly evident in rural labour markets as rural firms often face significant search costs in attracting and retaining workers with relevant skills, especially if the skills in demand are highly specialized (Rupasingha & Marré 2020).[6] To adjust for potential endogeneity resulting from labour market frictions, we apply the correction proposed by Ackerberg et al. (2015) (henceforth ACF). The main difference compared to LP is that ACF includes labour as an argument in the control function and that the first-stage estimation serves only to generate an estimate of $\phi_t'(k_{ijt}, m_{ijt}, l_{ijt})$ under the first-order Markov assumption (equation 2), which is then used to identify the input coefficients in the following second stage estimation:

$$y_{it} = \alpha + \gamma_l l_{ijt} + \gamma_k k_{ijt} + \gamma_m m_{ijt} + g(\phi_{t-1}' - \alpha - \gamma_k k_{ij,t-1} - \gamma_m m_{ij,t-1} - \gamma_{ijt} l_{ij,t-1}) + \xi_{ijt} + \eta_{ijt} \quad (7)$$

We use the following moment conditions to identify the coefficients for labor and capital:

$$\mathbb{E}\left[\phi_{ijt}^{'} \begin{pmatrix} k_{ijt} \\ l_{ijt-1} \end{pmatrix}\right] = 0. \quad (8)$$

and calculate TFP in the usual way as a residual using the estimates of the ACF production function.

A limitation with the ACF method is that it can be prone to bias in a gross-value version of the production function due to the scalar unobservable assumption (Ackerberg et al. 2015, Gandhi et al. 2020). That is, latent productivity is assumed to depend only on one scalar. In our empirical implementation, therefore, we use the approach proposed by Hu, Huang and Sasaki (2020) to evaluate the presence of measurement error in variable inputs.[7] Overall, we find that the ACF method is more consistent than those generated by alternative methods (OLS and LP), suggesting that it is important to account for labour market frictions in our institutional setting.

3.2 Additional worker qualities

One concern with the production function approach outlined above, is that there can exist complementarities between labour types and unobservable factors relating to additional worker qualities that are left unaccounted. There might also exist complementarities between firms internal and external access to knowledge that influence productivity. In our empirical application, therefore, we also estimate a specification where we relate a gross-value based

[6] In Sweden, employees are generally well protected against dismissal due to employment protection legislation. Such labour market regulations and search frictions are not distinctive features only for the Swedish context, but are found in labour markets and rural industries in several countries (Konings & Vanormelingen 2015, Donovan & Schoellman 2023).

[7] This implies including a least two intermediate inputs in the estimation. More information on this estimation and the intermediate inputs considered can be found in Section 5.

estimate of TFP to additional firm and worker qualities in the following by sub-industry:

$$TFP_{ijt} = \zeta_0 + \zeta_1 \, spec\_edu_{ijt} + \zeta_2 \, tenure_{ijt} + \zeta_{cj} C_{ijt} + r_{jt} + \tau + \upsilon_{ij} + x_{ijt} \quad (9)$$

where the first term denotes a vector of employment shares of skilled labour types and the second (tenure) measures occupation-specific labour market experience of all workers belonging to a firm. The associated productivity premiums are denoted ζ_1, ζ_2. In this estimation, we include county, time, and firm fixed effects denoted r_{jt}, τ_j and x_{ilt} is an error term. Included is also a vector of controls C_{ijt} measuring additional firm and worker qualities and regional conditions that can influence firms' possibility to source external knowledge. Via inclusion of firm fixed-effects, this estimation controls for informal skills not reflected in observables. The reduced form approach in equation 9 is similar to that used in Parrotta et al. (2014) and Konings and Vanormelingen (2015) who also regress gross-output based TFP on inputs that reflect the skill-level of employees.

3.3 Selection bias

Even with the inclusion of local labor market controls, fixed effects at the firm and county level, it is difficult to rule out that the choice of firms to employ skilled labour is endogenously determined. Firms could have sorted themselves into areas with a high supply of workers with the relevant skills, making it difficult to establish a causal link. Another concern is that results may be driven by less productive firms exiting the industry, causing selection issues related to panel attrition. We perform several robustness tests to address such selection issues. First, we employ a two-stage Instrumental Variable (IV) approach where we instrument the skill variables with indices that measure the pre-existing stock of labour with specialized training of relevance in the industry at the level of LA:s. More details on this estimation strategy and the validity of the instrument can be found in Section 5.3, and similarly computed labour supply instruments can be found in Parrotta et al. (2014) and Mohammadi et al. (2017). Second, and to account for selection issues related to firm exits, we re-estimate all models using the attrition correction proposed by Rovigatti and Mollisi (2018), which accounts for firms' probability of survival over the investigation period.

4 Data

The employer-employee matched data used in this study originate from several population registries governed by Statistics Sweden. The core of our data is information on individuals (employees and managers) from the Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA) merged with firm financial data from the business register.[8] Linking these registers, we obtain detailed information on all active firms, information on employees and managers of which the most important information for our purpose is firm financial accounts and knowledge characteristics of the workforce. We use five-digit Swedish Standard Industrial Classification codes (SNI) to distinguish firms with horse breeding and horse racing as the main source of economic activity. We have classified the industry according to the European Industry Classification codes (NACE, similar to NAICS for U.S and SNI for Sweden) in two main industries, horse breeding (SNI 1430) and horse racing (SNI 93191). Data from three additional registers are sourced to obtain information on firm size in terms of land holdings (ownership), land in production (including land in rental agreements) and firms geographical location; The Property Tax Register (FTR), the Swedish AICS/LPIS data (the Land Parcel Identification System) and the Geographical database (GDB).[9]

To provide time for the accumulation of skilled labour to take place within firms and for changes in productivity to materialize, we perform our production function estimation on data over a 13-year period (2010-2022). This period comprises several shocks including the 2018 drought that affected agrarian industries in Sweden adversely, the Covid-19 pandemic and its aftermath with steep input price inflation, which makes it highly relevant to apply the residual TFP method. The data set used in the analysis is an unbalanced panel of an average of 1,049 and 1,886 firms in horse breeding and racing, respectively, which amounts to 19,392 observations. We restrict the sample to exclude multi-plant firms, which amounts to about 1% of all firms, resulting in an estimation sample of 19,181 observations, that is, all single-plant firms with more than one employee (counting the manager) for which information on gross-value added is available.

Table 1 provides summary statistics of common input factors included in the model (estimation sample), most of which follow standard definitions in the literature. Table A4 in the Appendix provide detailed variable definitions and data sources.

[8]These micro data identify all firms, plants and individuals (over age 16) in the Swedish economy since 1990.

[9]We use data on the amount of pasture in ownership and rental agreement as an additional input in our production function estimation, see section 5 and Table A4 in the Appendix for the details.

Table 1: Summary statistics: production factors and skilled labour (workers include managers). Mean values 2010-2022 with standard deviations in parentheses.

	Industry		Horse breeding		Horse racing	
Gross-output	1518	(4454)	1680	(5027)	1436	(4122)
Intermediate inputs	1109	(4041)	1347	(4995)	978	(3240)
Capital	1346	(6871)	2222	(9589)	832	(4468)
Labour (FTE)	0.97	(3.32)	0.98	(3.88)	0.96	(2.93)
Land	3.48	(15.59)	6.38	(23.01)	1.78	(8.18)
Skilled labour						
Nr. workers spec. training (horse breeding, hippology)	0.13	(0.64)	0.11	(0.77)	0.14	(0.58)
Nr. workers spec. training (animal husbandry)	0.11	(0.45)	0.07	(0.37)	0.14	(0.50)
Nr. workers spec. training (agriculture)	0.11	(0.58)	0.15	(0.79)	0.10	(0.41)
Unskilled labour						
Nr. workers lack occupation-related training	2.05	(13.49)	1.67	(4.08)	2.16	(13.69)
Nr observations	19,181		7,158		12,116	
Nr firms	2,890		1,039		1,862	

Note: gross-output, intermediate inputs, and capital are displayed in thousand Swedish kronor KSEK (1 SEK approx. 0.090 USD) deflated using a producer input price index provided by the Swedish Board of Agriculture (2015=100). Labour is reported in Full Time Equivalent Employees (FTE).

4.1 Main variables: specialized training of relevance in the industry

The variables of main interest measure the skill level of the workforce and are constructed using information on educational qualifications among all workers belonging to a firm each year, counting both employees and managers. The first measures *occupation-specific* education defined as the number of workers with education in horse breeding and hippology constructed using 3-digit Standard Classification of Education (SUN) codes in the LISA register. SUN codes correspond to the International Standard Classification of Education (ISCED) and the 3-digit level is the most detailed including information about the specific subject for each individual. The remaining two variables measure *occupation-related* education defined as the number of workers with education in animal husbandry and agriculture (e.g. agronomy), respectively. The three educational qualifications are mutually exclusive (correlations are displayed in Table A1 in the Appendix). A key difference between these educational qualifications is that the former is most specific in terms of the type of training of relevance for a profession in the horse industry and perhaps for firm productivity. A person with an education in horse breeding and hippology is skilled in how to train horses and riders and they often work with horse care, feeding, and animal health issues. They also frequently act as advisors for horse farms and riding schools, and take part in research and development in the field (Asogwa et al. 2013). Our hypothesis is that workers with such formal qualifications should be able to support firms to build absorptive capacity and improved productivity via specialized skills, networks and external interactions. The remaining two education qualifications provide skills in agronomy, agricultural science and the care of other farm and domestic animals. Whether these related qualifications are productivity-enhancing in the horse industry depends on whether firms can benefit from having workers with complementary skills.

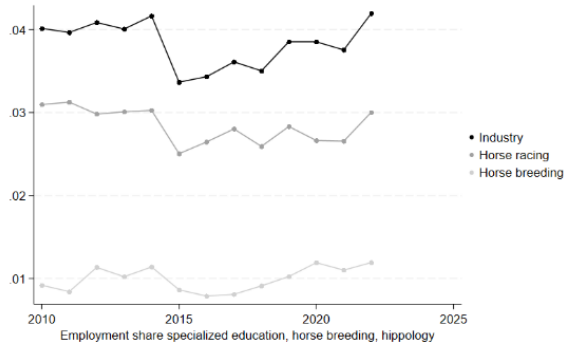
To further investigate the degree of complementarity in workers skills, we compute a variable that measures related knowledge diversity of a firm in the following:

$$H_{ijt}^{e} = -\sum_{e=1}^{r} p_{ijt}^{e} \ln p_{ijt}^{e} \tag{10}$$

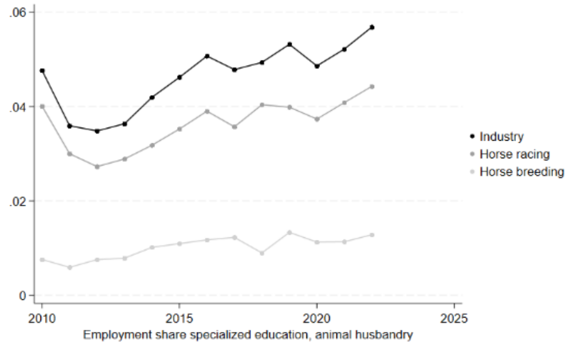
where p_{ijt}^{e} measures the workforce share in educational type e of firm i at time t, where e denotes the three educational qualifications defined above (horse breeding and hippology, animal husbandry and general agricultural education). Thus, we use both head count measures, employment shares and a diversity index to test hypotheses regarding the role of skilled labour for firm productivity in the industry.

Summary statistics (Table 1) show that the average firm in the industry has about one Full Time Equivalent (FTE) employee, indicating that firms are small on average, but there is considerable variation as evidenced by the large standard deviations of the amount of total labour and skilled labour employed in production. Figure 1 illustrates trends of the employment shares defined above over our 13-year period of investigation.[10] At industry level, the average employment share with specialized training in horse breeding and hippology is 4% with a standard deviation 16% and the share is increasing over time (Figure 1 panel a). The remaining employment shares and educational diversity are also rising over time and the standard deviations are large suggesting both variation and growth across firms. This shows that firms in the industry are up-skilling, but this is not uniform and there are large differences within the industry. All of the panels (a-d) show that the average firm specializing in horse racing has a workforce with more occupation-specific and occupation-related education then has an average firm in breeding. It is also evident that the average firm has increased its share of workers with both occupation-specific and occupational-related education qualifications over time and that this is present in both sectors of the industry.

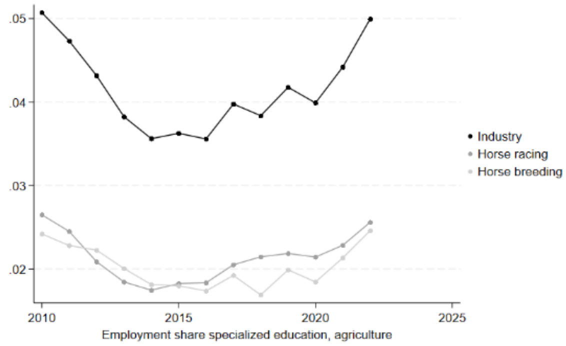
[10] Summary statistics and bivariate correlations regarding employments shares with occupations-specific and occupation-related education can be found in Table A2 and A3 in the Appendix.



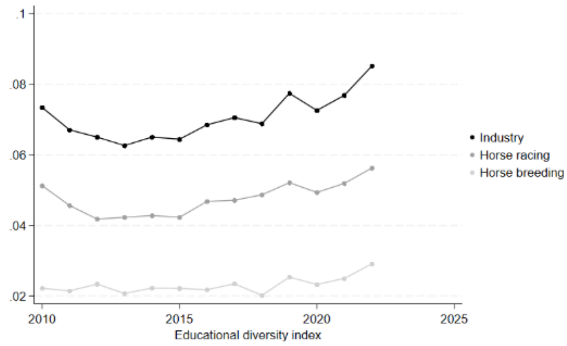
(a)



(b)



(c)



(d)

Figure 1: Employment share with specialized training 2010-2022

4.2 Firm, worker and labour market controls

The model includes controls for additional worker qualities commonly associated with firm productivity, summarized in Table 2. We include controls for education level, gender and age. The former is the most common proxy of firms' absorptive capacity in this literature (Fox & Smeets 2011) and the age of employees' controls for that younger individuals change jobs more often and receive more education and on-the-job training than older. We include controls for occupation-related experience to account for the mediating role played by tenure in the analysis of education (Lazear 2009). We compute this using information on individuals' labour market histories and measure the accumulated labour market experience in horse breeding or horse racing, respectively, held collectively by all workers and managers belonging to a firm each year. In calculating these variables, we use information on employment histories in the LISA register together with industry codes to track individuals' labour market experience (employment and self-employment) in horse breeding and horse racing from 1993 and onward.[11] The summary statistics (Table 2) show that an average firm specializing in horse breeding has an accumulated stock of experience in horse breeding of about 16 years and an accumulated stock of experience in horse racing of about 3 years. Accumulated experience is slightly higher among firms in horse racing (18 years in horse racing and 6 years in horse breeding), suggesting that there is movement of workers across firms.

The model includes a set of labour market controls measured at the municipality level, which is the most disaggregated unit of local governance in Sweden (there are 290 municipalities). We calculate two locational quotients that proxy for MAR externalities, such as access to specialized suppliers and intermediate input, which are frequently hypothesized external sources of firm productivity (Malmberg & Maskell 2002). These variables are computed by the number of firms in the two sectors (see Table A4). A variable that measures the supply of skilled labour in the municipality is included, defined as the stock of individuals with an education in horse breeding and hippology. This controls for firms located in areas with a high supply of workers with the relevant skills, which could be more likely to employ workers with such skills. Among regional controls, we also include a variable that measures the wage sum to account for regional size and density and to proxy firm access to thick markets and consumers with purchasing power (Jacobs 1969, Combes et al. 2012).[12] County-level fixed-effects are included to account for the time-invariant natural advantage and, lastly, we include controls for the legal status and landholdings of firms to account for additional moderating effects

[11] The LISA register has availability from 1990. Due to changes in Swedish industry codes (SNI) we cannot accurately track work experience by sub-industry before 1993. We have also calculated the tenure separately for self-employment (managerial experience) and employment, as in Braunerhjelm and Lappi (2023), but this distinction did not add to the analysis and we therefore excluded this distinction from the final model.

[12] An alternative measure of regional size commonly applied in the literature is overall population density. Since this is highly correlated with the wage sum, we use the latter due to its relevance for an experience based industry.

related to firm size (Medase 2020).

A limitation with the firm-level data is that they only report on firms established after 1986 (firms established before are given a value of 1986 regardless of establishment year). This prevents us from using a continuous measure of firm age as a quasi-fixed 'state' variable in the production function estimation (as in Manjón and Manez (2016)). Instead, we include age as a control that takes the value one if a firm is younger than 5 years. We also estimated separate models where we include the continuous measure of firm age together with a dummy variable denoting those firms for which age is unobservable prior to 1986 and the results are unaffected.

Table 2: Summary statistics. Additional worker qualities and controls. Mean values 2010-2022 with standard deviations in parentheses.

	Industry		Horse breeding		Horse racing	
Tenure horse breeding	15.55	(21.22)	16.57	(22.28)	6.32	(2.61)
Tenure horse racing	10.21	(17.81)	2.50	(4.09)	18.51	(57.24)
Share female	53	(15)	53.80	(14.27)	52.45	(15.38)
Average age employees, managers	0.97	(3.32)	0.98	(3.88)	0.96	(2.93)
Share higher education	0.01	(0.04)	0.01	(0.05)	0.01	(0.04)
Legal firm	0.17	(0.38)	0.17	(0.38)	0.18	(0.38)
Young firm	0.38	(0.40)	0.40	(0.44)	0.36	(0.49)
Location quotient, breeding	0.299	(0.320)	0.302	(0.344)	0.193	(0.204)
Location quotient, racing	0.399	(0.400)	0.398	(0.401)	0.401	(0.339)
Skilled labour (municipality)	126.05	(148.96)	124.06	(148.22)	127.06	(148.97)
Wage sum (municipality) /10000	15649	(35723)	14544	(34002)	15954	(36200)
Nr observations	19,181		7,158		12,116	
Nr firms	2,890		1,039		1,862	

Note: municipal wage sum is displayed in ten thousand Swedish kronor deflated using CPI (2015=100).

5 Regression results

5.1 Industry level

Table 3 displays the results from our production function estimations and Table 4 displays the results where we relate implied TFP to additional workers qualities measured as employment shares (equation 9). All production function estimations include a sub-industry control that equals one if a firm is specialized in horse breeding. From Table 3 results we can compare the estimates generated by the different productivity estimators, i.e. the standard methods (OLS, LP) and the ACF and evaluate the importance of contemporaneous input choice and labour market frictions in this industry. We report as well the results where we use the robust option proposed by Hu, Huang and Sasaki (2020) (hencefort HHS) to evaluate the presence of measurement error in intermediate inputs. Results show that labour and intermediate inputs are most important in explaining firm productivity, but the different approaches generate output elasticities that are different, qualitatively. In particular, OLS and LP are indicated to overestimate the output elasticity of intermediate and labour inputs and ACF and HHS are shown to generate more consistent estimtes as indicated by the lower standard errors. A first reflection is that results seem to support the theory that an upward bias in variable inputs is likely to occur when contemporaneous input choice and adjustment frictions are left unaccounted for (e.g. Levinsohn & Petrin 2003, Bond & Söderbom 2005, Ackerberg et al. 2007).[13] Figure 2 plots the kernel densities of estimated TFP using the different approaches, and overall these first estimations lend support to the view that labour is not a freely adjustable input. Furthermore, and since the ACF elasticities are indicated to be more precisely estimated, this is our preferred method in the productivity analysis that follows.

Next, we investigate how the introduction of quality differences between workers relate to productivity. Results (Table 3) indicate that skilled labour (measured as log headcounts of workers with education in either horse breeding and hippology, animal husbandry or agriculture) is associated with a higher productivity premium than is unskilled labour (the log of the number of workers that lack occupation-related education). Results also show that occupation-specific educational qualifications are associated with a higher elasticity ($0.331; p < 0.05$) than are occupation-related ($0.157, 0.139; p < 0.05$). These results support the hypothesis that firms should match workers of similar skills and education (Kremer 1993) and that it is specialized skills in horse breeding and hippology that matter most for firm productivity in the industry. Overall, we find very reasonable estimates of the gross-output production function while simultaneously correcting for transmission bias, labour market frictions and measurement errors. Intermediate inputs are associated with the highest elasticity, with an average of about 0.8 (ACF).[14] In Table 3, we also report a Wald test on the sum of the coefficients for each productivity estimator where the null is Constant Returns to Scale (CRS). All estimators, but the HHS, return a p-value where we can reject the null hypothesis, i.e. the gross-output production function for this industry does not exhibit CRS. The sum of the elasticities, a measure of returns to scale, is above 1 in all estimations (except HHS), suggesting that firms in the industry can scale inputs to increase efficiency.[15] The result that intermediate inputs are the predominant source of productivity growth, exceeding the contributions of both capital and labour, is consistent with studies focusing on related industries, such as agriculture (Skevas 2023).

[13] The differences between the estimates of labour and intermediate inputs are statistically significant across the different methods as evidenced by chi-square statistics and associated p-values, we can reject the null of equal coefficients at 5% confidence level.

[14] We have also tested educational diversity indices that measure overall diversity (across all educational subjects) and unrelated educational diversity (across all non-agricultural and non-horse related educational subject), but these estimates where always insignificant and was dropped from the final model.

[15] Since ACF estimates are associated with much lower standard errors compared to HHS and since the HHS estimate on CRS is significant at the 10% level, we believe that the CRS assumption can be rejected. Increasing returns to scale is also most consistent with previous studies of industries that exhibit similar conditions (see for example Skevas, 2023 on agriculture).

Table 3: Gross-output based estimates at industry level.

	Pooled sample, horse breeding and horse racing						
	OLS		LP		ACF		HHS
Log capital	0.009	0.018	0.072*	0.067*	0.025*	0.021*	0.022*
	(0.008)	(0.008)	(0.019)	(0.012)	(0.003)	(0.009)	(0.012)
Log labour	0.361*		0.326*		0.301*		
	(0.018)		(0.035)		(0.005)		
Log intermediate inputs	0.983*	0.983*	0.817*	0.866*	0.811*	0.801*	0.802*
	(0.007)	(0.007)	(0.026)	(0.023)	(0.005)	(0.005)	(0.009)
Log labour with spec. training		0.362*		0.341*		0.331*	0.330*
(horse breeding, hippology)		(0.047)		(0.043)		(0.002)	(0.007)
Log labour with spec. training		0.188*		0.186*		0.157*	0.158*
(animal husbandry)		(0.062)		(0.060)		(0.003)	(0.009)
Log labour with spec. training		0.171*		0.234*		0.139*	0.125*
(agriculture)		(0.058)		(0.055)		(0.002)	(0.007)
Log unskilled labour		-0.226*		-0.219*		-0.257*	-0.254*
		(0.037)		(0.039)		(0.002)	(0.009)
R-square	0.679	0.674					
Wald test CRS (p-value)			152.75	87.02	21446	35361	1.506
			(0.00)	(0.00)	(0.00)	(0.00)	(0.09)
Nr observations	19,181		19,181	19,181	19,181	19,181	19,181
Nr firms	2,890		2,890	2,890	2,890	2,890	

Note: LP and ACF are estimated with bootstrapped standard errors using 500 replications. OLS estimations are performed with robust standard errors clustered at firm-level based on the same estimation sample as LP and ACF. All estimations control for firm age. The second intermediate input considered in the HHS specification is the amount of pasture that is employed in production each year (not the amount in ownership or rental agreement). This additional input is sourced from the Land Parcel Identification System (LPIS) for Sweden, which can be linked to organization numbers in the data provided by Statistics Sweden.
* indicates statistical significance at the 5% level or lower.

Table 4 displays the results where we relate implied TFP generated by the ACF method to additional firm, workforce, and regional characteristics. Column 1 displays the results where we consider a between estimator, and the remaining columns show the results where we control for time-invariant firm-specific effects. In addition, the first two columns include workforce characteristics in terms of employment shares separating between the three educational types, and the last column displays the results where we include our Shannon entropy index of related educational diversity.

Results show that, when the educational variables are evaluated separately as shares and with the inclusion of firm-fixed effects (column 2), only the qualifications in horse breeding and hippology are positively related to productivity. In addition, when the model is estimated with the diversity index (column 3), there is a positive relationship between related educational diversity and productivity. These results may suggest two things. The first is that occupation-specific education is associated with an average productivity premium of about the size of 11% across firms in the industry. The second is that the contribution of educational diversity to productivity is most likely due to the combination of workers with specialized

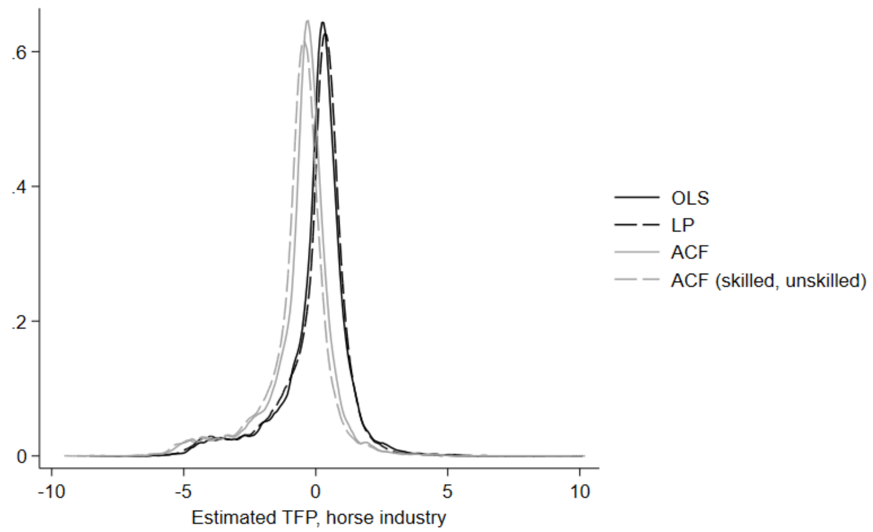


Figure 2: Kernel density of firm TFP at industry level calculated from different production function estimations with skilled and unskilled labour.

training in horses rather than the combination of workers with related skills. These industry-level results are again most in line with the hypothesis that firms in the horse industry should match workers of similar skills.[16]

The results (column 3) also show that occupation-specific education is associated with a higher productivity premium than experience in horse breeding and horse racing, respectively. This suggests that specific human capital acquired via formal training is relatively more important in explaining firm productivity than that acquired via labour market experience.[17] These results are quite consistent with previous studies based on employer-employee matched data showing positive and significant associations between human capital inputs and firm productivity, which are often found ranging between 1% and 17% in other industries (e.g. Pischke 2001, Frazis & Loewenstein 2005, Konings & Vanormelingen 2015)

A brief mention of the control variables shows that additional labor qualities measured as education level, seniority (average age of employees), and gender (share of female workers) are negatively associated with firm productivity, or are insignificant.[18] The coefficient that denotes the local clustering of firms in horse breeding is positive and significant. This could suggest that firms in the industry can benefit from access to specialized suppliers (breeders) and intermediate inputs that allow them to scale inputs to increase efficiency. Yet, this might also capture a combination of learning, sharing, and matching mechanisms of co-location or simply localized natural advantage that augment firm productivity. It is important to

[16]These differences are statistically significant at the 5% level.

[17]The null of equal coefficient can be rejected at 5% level (Chi-squre 120.21, $p < 0.05$).

[18]We have tested alternative educational diversity measures that account for overall diversity (across all educational types) and these are insignificant in all estimations.

stress that the specifications considered thus far are unable to disentangle the underlying mechanisms. A causal link between skilled labour and productivity is difficult to establish due to various sorting mechanisms. We return to this in Section 5.3 where we address the endogeneity of skilled labour in a causal effect analysis using an IV strategy.

Table 4: Productivity estimates at industry level with labour shares of occupational-specific educational qualifications. Pooled sample (breeding and racing)

Dependent variable: log TFP (ACF)	I (between)	II (between)	III (within)
Share spec. training horse breeding, hippology	0.126*	0.115*	-
	(0.052)	(0.009)	
Share spec. training animal husbandry	0.071	0.011	-
	(0.063)	(0.065)	
Share spec. training agriculture	0.223*	0.088	-
	(0.070)	(0.072)	
Educational diversity	-	-	0.171*
			(0.061)
Tenure, horse breeding	0.087*	0.033	0.035*
	(0.016)	(0.021)	(0.021)
Tenure, horse racing	0.073*	0.047*	0.048*
	(0.014)	(0.017)	(0.017)
Share higher education	0.009	0.010	0.010
	(0.010)	(0.022)	(0.021)
Share female	-0.153*	-0.154*	-0.150*
	(0.039)	(0.049)	(0.049)
Log average age employees, managers	-0.547*	-0.029	0.027
	(0.053)	(0.066)	(0.065)
Log land	0.077*	0.047*	0.047*
	(0.016)	(0.023)	(0.023)
Legal firm	0.130*	-	-
	(0.047)		
Location quotient, horse breeding	-0.062*	-0.087*	-0.087*
	(0.019)	(0.036)	(0.036)
Location quotient, horse racing	0.021	0.014	0.014
	(0.029)	(0.043)	(0.043)
Skilled labour (municipality)	0.027*	0.006	0.006
	(0.001)	(0.007)	(0.007)
Log wage sum (municipality)	-0.002	-0.004	-0.005
	(0.027)	(0.051)	(0.051)
Constant	1.503*	0.326*	0.312*
	(0.332)	(0.099)	(0.098)
R square	0.171	0.098	0.101
Nr observations	19,181	19,181	19,181
Nr firms	2,890	2,890	2,890

Note: robust standard errors in parenthesis clustered at firm level. All estimations include year and county dummies and a dummy to indicate sub-industry. * indicates statistical significance at the 5% level or lower.

5.2 Intra-industry heterogeneity

Table 5 and 6 shows the results where we use our estimation strategy to perform estimations among firms specializing in horse breeding and horse racing separately. This allows us to account for technology choice and obtain a more nuanced picture of the association between workforce characteristics and firm productivity within the industry. Table 5 results show that the intra-industry estimations are broadly in line with the industry-level results (Table 4), but with some notable differences regarding factor intensities. Both capital and skilled labour seems to be relatively more important inputs in horse racing compared to horse breeding, while intermediate inputs seems to be about equally important. This provides a first empirical verification of the common view that firms specializing in racing are more dependent on specialized know how than those specializing in horse breeding (McManus & Montoya 2012). Evaluations of the presence of simultaneity bias and adjustment frictions do also support a more nuanced picture than the industry level findings. The different approaches generate elasticities that differ in ways predicted by theory in that traditional estimates (OLS, LP) are indicated to overestimate the output elasticity of intermediate inputs (Figure 3). Furthermore, the ACF method is indicated to generate elasticities much larger than those estimated in the OLS and LP regressions, most notably in horse racing and w.r.t to workers with occupation-specific education. This suggests the existence of adjustment frictions and that firms cannot freely adjust their use of workers with specific training in horses. Thus, it appears that labour adjustment costs have differential influence in the two sectors. Results also suggest that specialized education in horses is associated with a much higher productivity premium in horse racing compared to horse breeding (this difference is statistically significant at the 5% level with a Chi-square statistic of 141.2 ($p < 0.05$).

Results from the extended productivity analysis (Table 6) show that we can confirm the important role played by labour market experience in breeding and racing for firm productivity in both sectors. However, occupation-related education is again only significant in horse breeding when we estimate the model across firms. In particular, when the models are estimated without firm-fixed effects, there is a positive relation between all three educational qualifications and productivity in horse racing (Column 4,5). The coefficient on agricultural education is the largest, which supports earlier evidence on the productivity-enhancing potential of agricultural schooling (Ali & Flinn 1989). When we evaluate instead the change that occurs within firms, it is only occupation-specific education that remains significant. Similarly to the above, the positive association between educational diversity and TFP can only be confirmed among firms specializing in horse racing (Column 6). These results could suggest that there exists significant within-firm variation that is correlated with firms' possibility to employ and sustain a skilled labour force. Firms that employ more skilled workers may simply have managers that are abler, but this may also relate to firm size. It could be

Table 5: Gross output estimates at intra-industry level with different control function approaches.

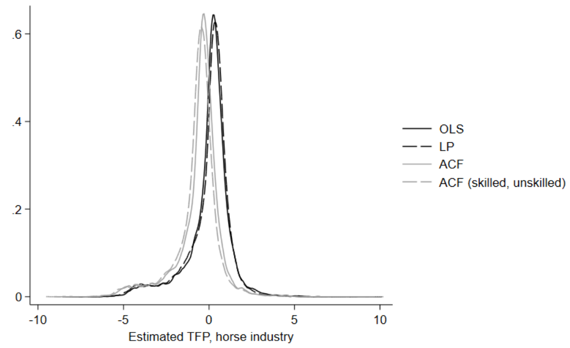
	Horse breeding				Horse racing			
	OLS	LP	ACF	HHS	OLS	LP	ACF	HHS
Capital	0.028*	0.010*	0.012*	0.012*	0.010*	0.078*	0.048*	0.049*
	(0.010)	(0.005)	(0.004)	(0.005)	(0.005)	(0.018)	(0.002)	(0.002)
Intermediate inputs	1.058*	0.837*	0.789*	0.788*	0.998*	0.854*	0.850*	0.850*
	(0.027)	(0.034)	(0.009)	(0.007)	(0.010)	(0.028)	(0.009)	(0.010)
Log labour with spec. training	0.255*	0.243*	0.103*	0.103*	0.477*	0.294*	0.456*	0.455*
(horse breeding, hippology)	(0.099)	(0.092)	(0.006)	(0.007)	(0.053)	(0.048)	(0.003)	(0.004)
Log labour with spec. training	0.254	0.049	0.029	0.022	0.281*	0.179*	0.275*	0.275*
(animal husbandry)	(0.223)	(0.125)	(0.031)	(0.032)	(0.064)	(0.062)	(0.003)	(0.007)
Log labour with spec. training	0.250*	0.247*	0.156*	0.155*	0.340*	0.267*	0.361*	0.360*
(agriculture)	(0.098)	(0.095)	(0.002)	(0.004)	(0.054)	(0.066)	(0.005)	(0.009)
Unskilled labour	0.073	0.082	0.057*	0.056*	-0.320*	-0.361*	-0.310*	-0.311*
	(0.070)	(0.074)	(0.006)	(0.005)	(0.041)	(0.049)	(0.003)	(0.005)
R-square	0.683			0.674				
Wald test CRS		29.02	2257.10	1899.79		34.44	52202	67023
(p-value)		(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)
Nr observations	7,158	7,158	7,158	7,158	12,116	12,116	12,116	12,116
Nr firms	1,039	1,039	1,039	1,039	1,862	1,862	1,862	1,862

Note: OLS estimations are performed with robust standard errors clustered at firm level, LP and ACF with bootstrapped standard errors using 500 replications. All estimations include controls for firm age.
* indicates statistical significance at the 5% level or lower.

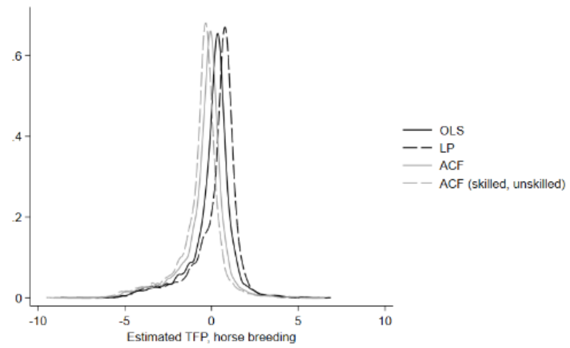
the case that both level and diversity in skills are more beneficial to larger firms as they have organizational routines, resources and financial strength that makes them better equipped to employ and sustain a workforce with specialized competence.

To investigate the moderating role of firm size, we reestimate the model including the size of firms' capital stock. The results are very similar, but the magnitude of the educational variables decreases slightly (see Table A3). We also evaluate if there is any change in the results when we consider only larger firms. Bearing in mind that firms in the industry are small, we have restricted possibilities to divide firms into multiple size groups. Therefore, we consider the subsample of firms that have a minimum of 3 FTEs (counting the manager(s)). The results (Table 7) show that the coefficient of occupation-specific educational diversity is larger in magnitude among larger firms (0.273) compared to the average across firms (0.215), but again only significant in horse racing. We can therefore conclude that firms specializing in horse racing are more likely to benefit from combinations of workers with specialized training in horse breeding and hippology and that this is likely to increase with firm size.

Taken together, we find robust evidence of a positive correlation between increased specialization in labor types and productivity among firms specializing in horse racing. These results could suggest that firms that invest in workers with specialized training in horses improve their productivity, perhaps via higher absorptive capacity and adoption of productivity-enhancing technologies and training methods and ways of doing business. This is a reasonable expla-



(a)



(b)

Figure 3: Kernel density of firm TFP at sub-industry level calculated from different production function estimations.

nation given the common interpretation of TFP as a proxy for technological progress (Solow 1957, Sala-i Martin & Barro 1995). Studies that focus on educated workers in an agrarian setting do also highlight that the main link between productivity and education is likely via technology adoption in that educated farmers are more willing to take risk and adopt new production technologies (Reimers & Klasen 2013). Below, we try to address the causal link between skilled labour and firm productivity, which makes results easier to interpret.

Another robust finding is that co-location can be important. This can be seen from the coefficient denoting local clustering of firms in horse breeding, which is positively associated with productivity among firms in racing. The reason why positive co-location might occur is that firms in racing can benefit from access to specialized horse suppliers and related inputs, which may lead them to increase efficiency. This could also be related to transport cost, which was Marshall's main explanation for the productivity advantages of co-location. Yet, this positive association may not only be attributed to the advantages of co-location, but can correlate with factors that we cannot observe in the registry data, such as inter-firm networks. Since the main focus of this study is on firms' access to internal knowledge, we do not attempt to disentangle the nature of this spillover.

Table 6: Productivity estimates at intra-industry level with labour shares of occupational-specific educational qualifications.

	Horse breeding			Horse racing		
Dependent variable: log TFP (ACF)	I	II	III	IV	V	VI
Share spec. training	0.140	0.166	-	0.121*	0.110*	-
(horse breeding, hippology)	(0.099)	(0.101)		(0.059)	(0.023)	
Share spec. training	0.015	0.025	-	0.153*	0.018	-
(animal husbandry)	(0.017)	(0.107)		(0.080)	(0.081)	
Share spec. training	0.039	0.078	-	0.418*	0.129	-
(agriculture)	(0.096)	(0.100)		(0.097)	(0.088)	
Educational diversity	-	-	0.098*	-	-	0.190*
			(0.047)			(0.041)
Tenure, horse racing	0.024	0.041	0.086	0.070*	0.015*	0.017*
	(0.058)	(0.049)	(0.075)	(0.016)	(0.009)	(0.009)
Tenure horse breeding	0.130*	0.084*	0.086*	0.026	0.010*	0.009
	(0.020)	(0.022)	(0.022)	(0.069)	(0.001)	(0.085)
Share higher education	0.003	0.009	0.007	0.004	0.003	0.003
	(0.004)	(0.011)	(0.008)	(0.005)	(0.004)	(0.004)
Share female	-0.173*	-0.208*	-0.207*	-0.164*	-0.118*	-0.111
	(0.060)	(0.082)	(0.082)	(0.049)	(0.061)	(0.061)
Log average age	-0.395*	-0.141	-0.140	-0.637*	0.164	0.158
(employees, managers)	(0.083)	(0.097)	(0.096)	(0.070)	(0.089)	(0.099)
Log land	0.074*	0.042	0.042	0.070*	0.047	0.048
	(0.019)	(0.027)	(0.027)	(0.027)	(0.037)	(0.037)
Legal firm	0.102	-	-	0.113*	-	-
	(0.079)			(0.057)		
Location quotient, horse breeding	-0.011*	-0.094*	-0.094*	0.032*	0.023*	0.020*
	(0.043)	(0.052)	(0.051)	(0.009)	(0.001)	(0.001)
Location quotient, horse racing	0.012	0.064	0.066	0.039	-0.003	-0.004
	(0.056)	(0.089)	(0.087)	(0.035)	(0.050)	(0.056)
Log skilled labour (municipality)	0.060	0.026	0.025	0.010*	0.049	0.037
	(0.055)	(0.088)	(0.087)	(0.001)	(0.044)	(0.065)
Log wage sum (municipality)	0.030*	0.109	-0.111	-0.007	0.038*	0.034*
	(0.003)	(0.089)	(0.089)	(0.035)	(0.009)	(0.009)
R square	0.101	0.077	0.078	0.162	0.076	0.079
Nr observations	7,158	7,158	7,158	12,116	12.116	12,116
Nr firms	1,039	1,039	1,039	1,862	1,862	1,858

Note: robust standard errors in parenthesis clustered at firm level. All estimations include year and county dummies. Specification I & IV (between), II, III, V & VI (within)
*indicates statistical significance at 5% level or lower.

5.3 Sensitivity analysis

Table 7: Productivity estimates at intra-industry level with labour shares of occupational-specific educational qualifications, firms with at least 3 Full Time Equivalent employees.

	$\geq$ 3FTE			
Dependent variable: log TFP (ACF)	Horse breeding		Horse racing	
	I(between)	II(within)	III(between)	IV(within)
Share spec. training	0.260		0.098*	
(horse breeding, hippology)	(0.202)		(0.009)	
Share spec. training	0.050		0.054	
(animal husbandry)	(0.097)		(0.049)	
Share spec. training	0.009		0.044**	
(agriculture)	(0.010)		(0.023)	
Educational diversity		0.232		0.202*
		(0.174)		(0.084)
Tenure horse breeding	0.050*	0.051*	0.043*	0.044*
	(0.010)	(0.010)	(0.007)	(0.007)
Tenure horse racing	0.089	0.088	0.037	0.035
	(0.063)	(0.062)	(0.029)	(0.027)
R square	0.121	0.121	0.139	0.144
Nr observations	1,459	1,459	4,401	4,401
Nr firms	295	295	702	702

Note: robust standard errors in parenthesis clustered at firm level. All estimations include year and county dummies and the firm, worker and labour market variables included in the main model.
*indicates statistical significance at 5% level or lower.

Several robustness checks are used to validate the main results. First, we employ a 2-stage least square model (2SLS) to estimate equation 9 for the sub-set of firms specializing in horse racing. In this estimation we instrument the Shannon index of educational diversity in the following first-stage estimation:

$$div\_edu_{it} = \zeta_0 + \zeta_1 \mathrm{LA}_{(i,k,t|1993)} + \zeta_2 \mathrm{tenure}_{it} + \zeta_{cj} C_{it} + \tau + r_{it} + x_{it} \tag{11}$$

where the instrument $\mathrm{LA}_{(i,k,t|1993)}$ is computed as the predicted stock of skilled labour in local labour market k at time t based on the stock observed in 1993. Specifically, it measures the stock of workers with occupation-related education in the active workforce computed w.r.t. educational qualifications in horse breeding and hippology.[19] Local labour markets (LA) are defined by Statistics Sweden based on observed commuting flows between municipalities, as such, they account for that the size of individuals' actual labour market and that workers can move beyond municipal borders. Our assumption is that pre-existing supply of skilled labour at the level of LA:s may be not correlated with a firm's current demand for skilled workers and productivity, if measured with a long time lag (in our case a minimum of 17 years and a maximum of 29 years). The instrument is significantly correlated with our index of educational diversity (17.9% $p < 0.001$) and is uncorrelated with TFP (2.1%, n.s.) conforming

[19]The predicted value is used in the estimation of equation 9 in place of the diversity index.

to the exclusion restrictions.[20] The F-statistics of the first stage is larger than 50 suggesting that the instrument is strong. The first and second stage results are presented in Table 8 and they are largely supportive of the main findings.

Table 8: First and second stage IV results for the sub-sample of firms specializing in horse racing.

Dependent variable (first stage): occupation-specific skills diversity H^{e}_{ijt}	Horse racing
Predicted stock of skilled labour in local labour market I^{h}_{iJc}	0.027*
	(0.003)
Constant	0.046*
	(0.003)
Firm, worker and labour market controls	Yes
F-value	52.75
R square	0.101
Nr observations	11,624
Nr firms	1,850
Dependent variable (second stage): log TFP (ACF)	
Educational diversity (predicted)	0.091*
	(0.004)
Constant	0.096*
	(0.021)
Firm, worker and labour market controls	Yes
Nr observations	11,624
Nr firms	1,850

Note: robust standard errors in parenthesis clustered at the LA level. All estimations include year and county dummies and the firm, worker and labour market variables included in the main model (Table 6), see Appendix for details.
* indicate statistical significance at the 5% level or lower.

As a second robustness check, we re-estimate all production functions using the ACF method with the attrition correction proposed by Rovigatti and Mollisi 2018. Results of the subsequent productivity analysis are displayed in Table A7 in the appendix showing that main results are very similar.

6 Conclusions

This study is related to a growing literature that focuses on workforce characteristics and firm productivity (e.g. Glaeser & Maré 2001, Fox & Smeets 2011, Almeida & Carneiro 2009, Konings & Vanormelingen 2015, De la Roca & Puga 2017, Serafinelli 2019, Crescenzi & Gagliardi 2018, Morris et al. 2020). We contribute to this literature with evidence from the

[20]Summary statistics regarding the instrument and additional information on how the national stock of workers with this type of specialized education has evolved over time can be found in the Appendix (Table A6).

equine industry that exemplifies a rural industry where specialized skills is a key part of the production process, but which is not formally classified as a knowledge intensive industry. Through our industry approach we hope to further broaden the debate on how different forms of skills and knowledge contribute to firm productivity in light of a growing knowledge-based rural economy (Naldi et al. 2015). Overall, there has been very few studies focusing on the role of human capital inputs on productivity among firms that have their main base in the rural economy and this is the first study focusing specifically on breeding and racing in the horse industry. Given the labour intensive nature of the work, we hypothesize that a high degree of occupation-specific training (measured as educational qualifications) can promote firm productivity in the industry.

With employer-employee matched data for all active firms and their employees (2010-2022) we can precisely measure educational qualifications and labour market experience of all workers and managers belonging to a firm. Thus, we expand previous research by focusing on occupation-specific training by including the entire workforce of a firm, rather than looking only at the managers (Sumner & Leiby 1987, Asadullah & Rahman 2009, Reimers & Klasen 2013, Nowak & Kijek 2016). We estimate a gross output production function to obtain a consistent estimate of firm TFP meaning that we do not have to resort to the unrealistic assumption that intermediate inputs are separable from primary inputs, such as labour and capital, which is usually not the case (Christensen 1975, Gandhi et al. 2020). Results show that firms in the industry can derive productivity gains from having workers with training specifically aimed at a profession in the industry, but this knowledge effect mainly arises among firms specializing in horse racing. Although our result are robust to a number of control function approaches that correct for simultaneity bias, labour market frictions and the endogeneity of skilled labour via IV estimations, this study is not without caveats. One limitation is that we lack data to examine other dimensions of firms absorptive capacity commonly considered in this literature, such as exports, R&D and network activities. Our data make it possible to observe firms that export, but this constitutes only a small fraction (less than 1% of all firms in the industry). Aware that we cannot fully capture the construct of absorptive capacity, we believe that our focus on the characteristics of the workforce is the most relevant given that labor is one of the most important inputs in the industry (McManus, 2014). We also do not have access to data to account for differences in the quality of education in our analysis, which would have been a useful extension. Therefore, our results may actually underestimate the influence of the studied educational qualifications on firm productivity.

The policy implications of our article are quite clear. The positive impact of specialized training in horse breeding and hippology on firm productivity supports the view that knowledge acquired vid formal training is a key factor that enhance productivity in the industry.

This is important information for firms and can support them to make informed labour investment decisions when it comes to matching workers with different skills. From a policy perspective, our results exemplify how to support productivity in an industry that is gaining interest in policy circles due to its significant contributions to rural and regional economies, its low-intensive nature and biodiversity enhancing use of permanents grasslands.

References

Ackerberg, D. A., Caves, K. & Frazer, G. (2015), 'Identification properties of recent production function estimators', *Econometrica* **83**, 2411–2451.

Ackerberg, D., Benkard, C. L., Berry, S. & Pakes, A. (2007), Econometric tools for analyzing market outcomes, *in* J. J. Heckman & E. E. Leamer, eds, 'Handbook of Econometrics', Vol. 6, North-Holland, Amsterdam, pp. 4171–276.

Ali, M. & Flinn, J. C. (1989), 'Profit efficiency among basmati rice producers in pakistan punjab', *American Journal of Agricultural Economics* **71**(2), 303–10.

Almeida, R. & Carneiro, P. (2009), 'The return to firm investments in human capital', *Labour Economics* **16**, 97–106.

Asadullah, M. N. & Rahman, S. (2009), 'Farm productivity and efficiency in rural bangladesh: the role of education revisited', *Applied Economics* **41**(1), 17–33.

Asogwa, V. C., Maduka, B. C. & Olaitan, S. O. (2013), 'Entrepreneurial skills required by instructors for training students in horse husbandry for job creation and wealth generation', *Journal of Emerging Trends in Educational Research and Policy Studies* **4**(4), 670–677.

Audretsch, D. B. & Belitski, M. (2023), 'Evaluating internal and external knowledge sources in firm innovation and productivity: an industry perspective', *RD Management* **53**(1), 168–192.

Barzotto, M. & De Propris, L. (2019), 'Skill up: Smart work, occupational mix and regional productivity', *Journal of Economic Geography* **19**(5), 1049–1075.

Becker, G. (1964), *Human Capital*, 3 edn, University of Chicago Press, The.

Black, S. E. & Lynch, L. M. (1996), 'Human-capital investments and productivity', *The American Economic Review* **86**(2), 263–267.

Boikos, S., Pinar, M. & Stengos, T. (2023), 'Bribery, on-the-job training, and firm performance', *Small Business Economics* **60**(1), 37–58.

Bond, S. & Söderbom, M. (2005), Adjustment costs and the identification of cobb douglas production functions (no, Technical report, 05/04). IFS Working Papers.

Boschma, R., Eriksson, R. & Lindgren, U. (2009), 'How does labour mobility affect the performance of plants? the importance of relatedness and geographical proximity', *Journal of Economic Geography* **9**(2), 169–190.

Braunerhjelm, P. & Lappi, E. (2023), 'Employees' entrepreneurial human capital and firm performance', *Research Policy* **52**(2), 104703.

Christensen, L. R. (1975), 'Concepts and measurement of agricultural productivity', *American Journal of Agricultural Economics* **57**(5), 910–915.

Combes, P. P., Duranton, G., Gobillon, L., Puga, D. & Roux, S. (2012), 'The productivity advantages of large cities: Distinguishing agglomeration from firm selection', *Econometrica* **80**(6), 2543–2594.

Crescenzi, R. & Gagliardi, L. (2018), 'The innovative performance of firms in heterogeneous environments: The interplay between external knowledge and internal absorptive capacities', *Research Policy* **47**(4), 782–795.

Cunningham, C., Foster, L., Grim, C., Haltiwanger, J., Pabilonia, S. W., Stewart, J. & Wolf, Z. (2023), 'Dispersion in dispersion: Measuring establishment?level differences in productivity', *Review of Income and Wealth* **69**(4), 999–1032.

De la Roca, J. & Puga, D. (2017), 'Learning by working in big cities', *Review of Economic Studies* **84**(1), 106–42.

Donovan, K. & Schoellman, T. (2023), 'The role of labor market frictions in structural transformation', *Oxford Development Studies* **51**(4), 362–374.

Fox, J. T. & Smeets, V. (2011), 'Does input quality drive measured differences in firm productivity?', *International Economic Review* **52**(4), 961–989.

Frazis, H. & Loewenstein, M. A. (2005), 'Reexamining the returns to training: Functional form, magnitude, and interpretation', *Journal of Human Resources* **40**(2), 453–476.

Gandhi, A., Navarro, S. & Rivers, D. A. (2020), 'On the identification of gross output production functions', *Journal of Political Economy* **128**(8), 2973–3016.

Glaeser, E. L. & Maré, D. C. (2001), 'Cities and skills', *Journal of Labor Economics* **19**(2), 316–342.

Heldt, T., Macuchova, Z. & Alnyme, O. (2018), 'National and regional economic effects of the horse industry in sweden estimations from a bi-model for 2016: Estimations from a bi-model for 2016', *Human-horse relationships in work and play, Leeds* pp. 19–21.

Hu, Y., Huang, G. & Sasaki, Y. (2020), 'Estimating production functions with robustness against errors in the proxy variables', *Journal of Econometrics* **215**(2), 375–398.

Iranzo, S., Schivardi, F. & Tosetti, E. (2008), 'Skill dispersion and firm productivity: An analysis with employer-employee matched data', *Journal of Labor Economics* **26**(2), 247–285.

Jacobs, J. (1969), *The economy of cities*, Random House, New York.

Jez, C., Coudurier, B., Cressent, M. & Méa, F. (2013), 'Factors driving change in the french horse industry to 2030', *Advances in Animal Biosciences* **4**(2), 66–105.

Konings, J. & Vanormelingen, S. (2015), 'The impact of training on productivity and wages: firm- level evidence', *Review of Economics and Statistics* **97**(2), 485–497.

Kremer, M. (1993), 'The o-ring theory of economic development', *The Quarterly Journal of Economics* **108**(3), 551–575.

Lazear, E. P. (2009), 'Firm-specific human capital: A skill-weights approach', *Journal of Political Economy* **117**(5), 914–940.

Levinsohn, J. & Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *The Review of Economic Studies* **70**(2), 317–341.

López, C. L., Celaya, R. F., García, L. M. M., Rodrigues, U., M., M. A. & Osoro, K. (2019), 'Comparative foraging behaviour and performance between cattleand horses grazing in heathlands with different proportions of improved pasture area', *Journal of Applied Animal Research* **47**, 377–385.

Malerba, F. (1992), 'Learning by firms and incremental technical change', *The Economic Journal* **102**(413), 845–859.

Malmberg, A. & Maskell, P. (2002), 'The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering', *Environment and Planning A: Economy and Space* **34**(3), 429–449.

Manjón, M. & Manez, J. (2016), 'Production function estimation in stata using the ackerberg–caves– frazer method', *The Stata Journal* **16**(4), 900–916.

Marschak, J. & Andrews, W. (1944), 'Random simultaneous equations and the theory of production', *Econometrica* **12**, 143–205.

McManus, P. & Montoya, D. (2012), 'Toward new understandings of human–animal relationships in sport: A study of australian jumps racing', *Social Cultural Geography* **13**(4), 399–420.

Medase, S. K. (2020), 'Product innovation and employees' slack time. the moderating role of firm age size', *Journal of Innovation Knowledge* **5**(3), 151–174.

Mohammadi, A., Broström, A. & Franzoni, C. (2017), 'Workforce composition and innovation: how diversity in employees' ethnic and educational backgrounds facilitates firm?level innovativeness', *Journal of Product Innovation Management* **34**(4), 406–426.

Morris, D., Vanino, E. & Corradini, C. (2020), 'Effect of regional skill gaps and skill shortages on firm productivity', *Environment and Planning A: Economy and Space* **52**(5), 933–952.

Naldi, L., Nilsson, P., Westlund, H. & Wixe, S. (2015), 'What is smart rural development?', *Journal of Rural Studies* **40**, 90–101.

Noteboom, B. (2000), *Learning and Innovation in Organizations and Economies*, Oxford University Press, Oxford.

Nowak, A. & Kijek, T. (2016), 'The effect of human capital on labour productivity of farms in poland', *Studies in Agricultural Economics* **118**(1), 16–21.

Olley, G. S. & Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**, 1263–1297.

Orr, S. (2022), 'Within-firm productivity dispersion: Estimates and implications', *Journal of Political Economy* **130**(11), 2771–2828.

Parrotta, P., Pozzoli, D. & Pytlikova, M. (2014), 'Labor diversity and firm productivity', *European Economic Review* **66**, 144–179.

Penrose, E. T. (2009), *The Theory of the Growth of the Firm*, Oxford university press.

Pickering, C. M. H., Newsome, W., Leung, D. & F., Y. (2009), 'Comparing hiking, mountain biking and horse-riding impacts on vegetationand soils in australia and the united states of america', *Journal of Environmental Management* **91**, 551–562.

Pischke, J. S. (2001), 'Continuous training in germany', *Journal of Population Economics* **14**, 523–548.

Reimers, M. & Klasen, S. (2013), 'Revisiting the role of education for agricultural productivity', *American Journal of Agricultural Economics* **95**(1), 131–152.

Rigby, D. L. & Essletzbichler, J. (2006), 'Technological variety, technological change and a geography of production techniques', *Journal of Economic Geography* **6**(1), 45–70.

Rovigatti, G. & Mollisi, V. (2018), 'Theory and practice of total-factor productivity estimation: The control function approach using stata', *The Stata Journal* **18**(3), 618–662.

Rupasingha, A. & Marré, A. W. (2020), 'Moving to the hinterlands: Agglomeration, search costs and urban to rural business migration', *Journal of Economic Geography* **20**(1), 123–153.

Sala-i Martin, X. X. & Barro, R. J. (1995), Technological diffusion, convergence, and growth (no, Technical report, 735). Center discussion paper.

Serafinelli, M. (2019), '"good" firms, worker flows, and local productivity', *Journal of Labor Economics* **37**(3), 747–792.

Skevas, I. (2023), 'A novel modeling framework for quantifying spatial spillovers on total factor productivity growth and its components', *American Journal of Agricultural Economics* **105**(4), 1221–1247.

Solow, R. M. (1957), 'Technical change and the aggregate production function', *The review of Economics and Statistics* **39**(3), 312–320.

Sumner, D. A. & Leiby, J. D. (1987), 'An econometric analysis of the effects of human capital on size and growth among dairy farms', *American Journal of Agricultural Economics* **69**(2), 465–470.

Wixe, S. & Andersson, M. (2017), 'Which types of relatedness matter in regional growth? industry, occupation and education', *Regional Studies* **51**(4), 523–536.

Zasada, I. B., Hilgendorf, R. & J. Piorr, A. (2011), 'Horsekeeping and the peri-urban development in the berlin metropolitan region', *Journal of Land Use Science* **8**, 199–214.

Östbring, L. & Lindgren, U. (2013), *Labour mobility and plant performance*, Geografiska Annaler: Series B, Human Geography, 95(4), 287-305.
stergaard et al.

Østergaard, C. R., Timmermans, B. & Kristinsson, K. (2011), 'Does a different view create something new? the effect of employee diversity on innovation', *Research Policy* **40**(3), 500–509.

Appendix

Table A1: Correlation matrix. Skilled and unskilled labour measured as headcounts.

	1.	2.	3.
1. Workers with specialized training in horse breeding and hippology			
2. Workers with specialized training in animal husbandry	0.333		
3. Workers with specialized training in agriculture	0.540	0.261	
4. Unskilled workers	0.174	0.154	0.302

Table A2: Summary statistics. Additional worker qualities and controls. Average values 2010-2022. Mean values and standard deviations in parentheses.

	Industry		Horse breeding		Horse racing	
Share specialized training horse breeding, hippology	0.04	(0.16)	0.03	(0.13)	0.05	(0.17)
Share specialized training animal husbandry	0.05	(0.18)	0.02	(0.14)	0.06	(0.21)
Share specialized training agriculture	0.04	(0.18)	0.04	(0.18)	0.03	(0.15)
Educational diversity	0.06	(0.18)	0.05	(0.17)	0.08	(0.19)
Nr observations	19,181		7,158		12,116	
Nr firms	2,890		1,039		1,862	

Note: municipal wage sum is displayed in ten thousand Swedish kronor KSEK (1 SEK 0.090 USD) deflated using a consumer price index provided by Statistics Sweden (2015=100).

Table A3: Correlation matrix. Skilled unskilled labour measured as employment shares.

	1.	2.	3.
1. Share specialized training horse breeding and hippology			
2. Share specialized training animal husbandry	-0.026		
3. Share specialized training agriculture	-0.033	-0.035	
4. Educational diversity index	0.479	0.590	0.546

Table A4: Detailed variable definitions.

Variable	Definition
Gross output	Total sales in KSEK.
Intermediate inputs	Value of intermediate inputs in KSEK.
Labour	Number of Full Time Equivalent employees including manager(s).
Capital	Value of material and immaterial assets (machinery, buildings, and land) in KSEK.
Share specialized training, horse breeding, hippology	Share of workforce with education in horse breeding and animal husbandry constructed using Standard Classification of Education (SUN) codes (621f).
Share specialized training, animal husbandry	Share of workforce with education in animal husbandry (SUN 621e, 621g).
Share specialized training, agriculture	Share of workforce with education in agriculture, such as agronomy (SUN 620x, 621a-621d, 621x).
Educational diversity	Equation 10.
Tenure breeding	Accumulated years of experience (employment and self-employment) in horse breeding among workers (incl. manager(s)) measured from 1993.
Tenure racing	Accumulated years of experience (employment and self-employment) in horse racing among workers (incl. manager(s)) measured from 1993.
Land	Number of hectares of agricultural land (pasture, arable) in ownership or rental agreement.
Young firm	Equals one if less than 5 years since establishment.
Legal status	Equals one if the firm is registered as a legal company. The base is sole proprietorships and trading companies.
Share female	Share of employees (incl. manager(s)) that are female.
Share higher education	Employment share (incl. manager(s)) with a bachelor degree or above regardless of type of education.
Average age employees	Average age of employees (incl. manager(s)).
Locational quotient, horse breeding	Location quotient: The share of firms in horse breeding in municipality divided by the share of firms in horse breeding in the national total. Excluding the focal firm.
Locational quotient, horse racing	Location quotient: The share of firms in horse racing in municipality divided by the share of firms in horse racing in the national total. Excluding the focal firm.
Skilled labour (municipality)	The stock of individuals in local labour market (LA) with an education in horse breeding and hippology.
Wage sum (municipality)	Real wage sum in municipality measured in 10,000 KSEK.

Table A5: Productivity estimates at intra-industry level with labour shares of occupational-specific educational qualifications. Including a control for firms' capital stock.

	Horse breeding		Horse racing	
Dependent variable: log TFP (ACF)	Within		Within	
Share specialized training horse breeding, hippology	0.166		0.109***	
	(0.101)		(0.022)	
Share specialized training, animal husbandry	0.025		0.013	
	(0.107)		(0.080)	
Share specialized training, agriculture	0.078		0.126	
	(0.100)		(0.084)	
Educational diversity		0.097		0.188***
		(0.077)		(0.040)
Tenure, horse racing	0.041	0.085	0.014**	0.017***
	(0.049)	(0.074)	(0.009)	(0.008)
Tenure horse breeding	0.084***	0.085***	0.010*	0.005
	(0.022)	(0.022)	(0.001)	(0.076)
Share higher education	0.009	0.004	0.002	0.002
	(0.011)	(0.008)	(0.004)	(0.004)
Share female	-0.208***	-0.204***	-0.116**	-0.114*
	(0.082)	(0.080)	(0.060)	(0.061)
Average age employees	-0.141	-0.141	0.165*	0.159*
	(0.097)	(0.094)	(0.088)	(0.088)
Land	0.042	0.040	0.044	0.045
	(0.027)	(0.025)	(0.037)	(0.037)
Capital	0.022***	0.021***	0.030***	0.031***
	(0.007)	(0.006)	(0.008)	(0.008)
Labour market controls	Yes	Yes	Yes	Yes
R square	0.077	0.079	0.077	0.080
Nr observations	7,158	7,158	12.116	12,116
Nr firms	1,039	1,039	1,862	1,858

Note: robust standard errors in parenthesis clustered at firm level. All estimations include year and county dummies and the municipality level controls included in the baseline model (Table 6).
*** $p < 0.01, **p < 0.05, *p < 0.10$

Table A6: Table A6. Summary statistics IV. Mean values and standard deviations in parentheses.

	Horse racing	
Index skilled labour (LA)	1.057	(0.64)
Skilled labour in municipality 1990	29.96	(26.76)
Skilled labour in municipality at t	126.92	(148.90)
Skilled labour in LA 1990	281.92	(311.03)
Skilled labour in LA at t	1,290.83	(1,523.19)

Table A7: Productivity estimates at intra-industry level with labour shares of occupational-specific educational qualifications.

	Horse breeding		Horse racing	
Dependent variable:				
(ACF with the attrition option in Rovigatti and Mollisi 2018)				
Share specialized training, horse breeding, hippology	0.163		0.112*	
	(0.103)		(0.022)	
Share specialized training, animal husbandry	0.026		0.017	
	(0.108)		(0.084)	
Share specialized training, agriculture	0.079		0.123	
	(0.102)		(0.089)	
Educational diversity		0.099*		0.191*
		(0.044)		(0.040)
R square	0.079	0.081		0.077
0.082				
Nr observations	7,158	7,158	12.116	12,116
Nr firms	1,039	1,039	1,862	1,858

Note: robust standard errors in parenthesis clustered at firm level. All estimations include year and county dummies and the firm, worker and labour market variables included in the main model (Table 6).
* denote significance at the 5% level or lower.