



Training and productivity in micro-firms

Marco Rispoli ^{a,1}, Federico Tamagni ^b ^{*}

^a Practice SIT Research Center, Prometeia, Bologna, Italy

^b Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy

ARTICLE INFO

JEL classification:

J24
D22
C21

Keywords:

Micro-firms
Training
Productivity
IPWRA

ABSTRACT

This paper examines the effect of training on productivity of micro-firms, defined in official statistics as firms with less than 10 employees. While training may represent a crucial mode of knowledge accumulation in micro-firms, which usually lack resources or structure to devote to formal innovation activities, the specificities of micro-businesses suggest that training may even be detrimental to productivity of this category of firms. Yet, the received wisdom that training should foster their productivity, as it does in larger firms, remains untested, mainly due to difficulties in previous literature to analyze micro-firms separately from the wider groups of small or small-medium enterprises. Exploiting a unique administrative panel dataset that records training expenses of a large sample of micro-firms located in three Italian regions, we address whether micro-firms which train enjoy a productivity premium vis-a-vis micro-firms that do not. Our main finding, controlling for endogeneity of training via Inverse Probability Weighted Regression Adjustment methods, combined with firm fixed-effects, is a positive and sizable average treatment effect of training, stable across a series of robustness checks and heterogeneity analysis.

1. Introduction

In a globalized and increasingly digitalized economy, the human capital, skills and knowledge of employees are considered key determinants of firm competitiveness. As one of the main modes of accumulation of knowledge and human capital within companies, the provision of training, is often seen as a crucial strategy to promote firm innovation, performance and ultimately productivity. However, training remains a rare practice among firms of relatively small size, particularly the group of very small, so-called micro-firms, defined by official statistics as firms with less than 10 employees. This raises worries particularly in Europe, as the lack of skilled workers and poor knowledge upgrading characterizes about one third of European SMEs (Farvaque et al., 2009), a key reason being often identified exactly in the historically low provision of training in micro-firms (Reid, 1993; CEDEFOP, 2015). In fact, there is a widespread tendency to consider micro-firms as marginal entities with no innovation capabilities and scarce dynamism. This in spite of the existing empirical evidence supporting that proactive and dynamic firms, capable of meaningful innovation activity, do exist even within micro-firms (Brancati et al., 2022) and that micro-firms – in general, not just the very small high-tech start-up companies – are able to benefit from innovation in ways comparable to larger SMEs (Baumann and Kritikos, 2016). This literature typically measures innovation directly, via R&D, patenting,

or product/process innovation indicators, and asks whether such formalised innovation types translates into performance gains. Our study provides complementary evidence by focusing on training, arguably a central mode of knowledge accumulation available to micro-firms, whose innovation and performance dynamics are, by their very nature, centered around knowledge and competencies of their core people.

Several explanations have been proposed to account for the observed under-investment (or zero investment) in training by firms of very small size. Within standard conceptualizations of human capital accumulation that look at incentives of profit-maximizing rational firms (Becker, 1964; Acemoglu and Pischke, 1998, 1999), the “risk of poaching hypothesis” suggests that small firms may prefer not to invest in training since their trained employees are very likely to leave the firm attracted by better job opportunities offered by larger companies, while the “ignorance explanation” and the “market-forces explanation” (Westhead and Storey, 1996), respectively point to myopic behavior of small-firms and significant market failures. Other theories, mostly developed within the literature that applies the resource-based or capability-based view of the firm to the context of small-medium firms (see e.g. Kelliher and Reinl, 2009; Yazdanfar, 2013; Sheehan, 2014), offer alternative explanations. These frameworks stress that low (or zero) investment in training may be in fact an appropriate choice by smaller firms in line with their specificity,

* Correspondence to: Institute of Economics, Scuola Superiore Sant'Anna, Piazza Martiri 33, 56127, Pisa, Italy.

E-mail addresses: marco.rispoli@prometeia.com (M. Rispoli), federico.tamagni@santannapisa.it (F. Tamagni).

¹ Disclaimer: The views and opinions expressed here do not reflect the views or positions of Prometeia SPA.

unrelated to sub-optimal response to incentives or external factors. The required competencies and knowledge are multifaceted, while creativity and flexibility are needed. The typical training packages may not serve the needs of smaller businesses as they may result into very specialized staff, perhaps too much, so that learning and knowledge accumulation are better coped via on-the-job training, self-directed learning or coaching (Devins and Johnson, 2003; Kotey and Slade, 2005; Lopez et al., 2019).

Whatever the explanation for low provision of training, in this paper we address the even more basic question whether training does indeed positively contribute to productivity in very small, micro-firms. In fact, the different theories about why small firms may not invest in training, hinder a deeper disagreement on this even more basic question. The ignorance, market failure and poaching risk explanations, by relying upon a standard neoclassical production function conceptualization of the relations linking human capital and knowledge accumulation to firm performance, all postulate a sort of automatic positive effect of training on productivity. Conversely, the existence of a productivity-premium of training is questioned or even negated in the resource-based or capability-based view of the firm, incorporating the conceptualization of knowledge accumulation, learning and innovation as evolutionary processes (Dosi et al., 2000), occurring collectively at the level of the firm as an organization, not at the level of individual employees. In these approaches training is not just about upgrading of task-specific skills of employees, but it rather represents a broader mode of organizational knowledge accumulation. However, the channel going from expanding knowledge via training to productivity, is not working automatically. The outcome – not only in the case of firms of smaller size, but more generally in firms of any size – is heavily dependent on firm-specific characteristics and knowledge base, resting on idiosyncratic routines and heuristics. In fact, investing in training in itself can be also seen as a structural characteristic, distinctive of firms that are more proactive toward knowledge accumulation and innovation, but whose effectiveness in helping to climb the productivity ladder is contingent to firm-specific heterogeneity (Dosi et al., 2021).

All in all, considering these contrasting theoretical predictions, whether training is beneficial to productivity in micro-firms ultimately remains an empirical question. But there are no empirical studies – to our knowledge – of the training–productivity link in micro-firms. This research gap reflects the general tendency of the literature to under-research micro-firms as entities that merit separate attention from small or medium enterprises. From an empirical standpoint, this may be partly explained by the traditionally limited availability of large and detailed data on knowledge dynamics in micro-firms (Reid, 2007). But limited theory development also plays a role, since the existing conceptual frameworks – even the theories that we mentioned above – tend to lack a specific reference to micro-businesses. In spite of the increasing awareness that SMEs or even small firms are too broad categorizations, mixing together quite heterogeneous organizations (Gherhes et al., 2016), and notwithstanding the multiple calls for devoting specific research focus to micro-firms (Sels et al., 2006; Lai et al., 2017; Wright et al., 2015), studies which separate micro-firms from the wider groups of small-firms or SMEs remain scant. Regarding the context of our study, the state of the art is particularly unsatisfactory, since information on training investment by micro-firms is rarely available, usually only for the limited samples covered in ad-hoc surveys or single-firm case studies. The few existing empirical studies which do have information on training and do separate micro-firms from larger entities, either (i) explore the effect of training (or knowledge acquisition) on dimensions of performance other than productivity; or (ii) take productivity as the outcome variable of interest but do not isolate the effect of training (or knowledge acquisition) from that of other aspects of firms' human resource development strategy.

Beyond addressing the basic question whether training has a beneficial effect on productivity of micro-firms, we also explore the variation

of the estimated effect of training across four dimensions of heterogeneity, namely firm size, extent of firms' investment in tangible and intangible assets, and the specificity of manufacturing vis-a-vis other sectors. We explicitly test for possible complementarity (or substitutability) between training and these heterogeneity dimensions, allowing us to provide at least suggestive evidence about some possible channels shaping the productivity-training nexus in micro-firms.

Our study can also have relevant policy implications. In fact, an unintended and undesirable consequence of the limited attention devoted to micro-firms, is that existing policies targeting knowledge accumulation in SMEs often fail to recognize micro-firms as deserving a specific, separate intervention (Kotey and Slade, 2005; Wright et al., 2015). Were training to turn out as beneficial to productivity of micro-firms from our estimates, then policymakers could hope to succeed in boosting productivity of this group of firms through interventions designed to encouraging training investment in micro-firms or removing the obstacles that hamper training investment by them. This is an attractive area for broadly intended innovation policies aimed at closing the knowledge and productivity gaps of micro-firms, especially in Europe, since micro-firms represent between 80% and 95% of the firms active in most countries (see, e.g. CEDEFOP, 2015; OECD, 2017).

The empirical analysis exploits access to a unique dataset, developed within a joint project between the Italian Statistical Office (ISTAT) and one of the main associations of Italian SMEs (Confederazione Nazionale dell'Artigianato e Piccola-Media impresa, CNA), allowing to track training expenses in a large representative sample of micro-firms active in three Italian regions over the period 2013–2017. Italy represents a particularly interesting testbed, as the country has one of the highest shares of micro-firms among European economies, in all sectors of activity (Hall et al., 2009; Mitri et al., 2013). Moreover, in Italy, micro-firms show a particularly marked productivity gap vis-a-vis other firms, wider than in other countries (according to data from the OECD-MultiProd project, see Berlingieri et al., 2017). This connects to the debate about the so-called “neo-dualism” in Italian productivity dynamics, characterized by a persistent division between a small set of high-productivity firms and a large group of low-productive laggards (Dosi et al., 2012). This divergence has been related (Caravella et al., 2021) to ample heterogeneities in the innovation strategies of Italian firms, across more dynamic/knowledge intense firms vs. low innovative/cost competing firms. There is evidence (Brancati et al., 2022) that dynamic/proactive firms do exist even within Italian micro-firms, but there is no direct analysis of whether knowledge accumulation related to training improves their performance. Existing studies on the training–productivity link in Italy (Mancinelli and Mazzanti, 2009; Dosi et al., 2021) – much like the international literature – do not devote specific focus to micro-firms. Our paper can provide a contribution also in this respect.

As repeatedly acknowledged in the literature (Black and Lynch, 1996; Dearden et al., 2006; Zwick, 2006; Konings and Vanormelingen, 2015), estimating the effect of training on firm performance (productivity in our context) requires to account for endogeneity of training investment. In absence of sources of exogenous variation, in this study we couple firm fixed-effects with Inverse Probability Weighted Regression Adjustment (IPWRA), a doubly robust methodology allowing to identify Average Treatment Effect (ATE) under endogenous selection into the group of training firms (Słoczyński et al., 2022). In this respect, this paper is also an attempt to respond to the long-lasting “lost quest” in the micro-business literature for improved identification of the relationship between training and micro-firm performance (Patton et al., 2000). We indeed hope to contribute, more generally, to raise awareness that advancements in this direction are especially needed in the literature on small businesses (Bryan, 2006), particularly to progressing in the design of evidence-based policies (Devins and Johnson, 2003).

Our main finding, corroborated by covariate-balance analysis in the IPWRA, inclusion of firm fixed-effects and a series of robustness analysis, is that training significantly and positively affects productivity. As

the control group is made of micro-firms which do not train, our results support that training has an additional productivity-enhancing effect, on top of the other informal modes of knowledge accumulation and competence upgrading, which are presumably in place in both training and non-training micro-firms.

The heterogeneity analysis reveals that the positive contribution of training to productivity of micro-firms, can be seen as general-purpose, not tightly dependent from any of the dimensions of heterogeneity considered. However, investing in training – while neutral to sector membership and to co-occurrence of intangible investments – appears as possibly more effective, comparatively, in fostering productivity in single-employee (the sole owner/manager) firms and in firms which, by not investing in tangibles, more heavily rely upon their workforce as a source of knowledge accumulation and competitiveness.

2. Background literature and working hypotheses

2.1. Theories of training and productivity: the specificity of micro-firms

From a theoretical viewpoint, the relationship between training and productivity is commonly understood through the lenses of conventional neoclassical human capital theory, within a standard production function framework. This literature has evolved around the debate on the incentives that firms may have in paying the cost of training for their employees. In the original [Becker \(1964\)](#)'s model with perfectly competitive labor markets, firms have an incentive to invest in specific training, but not in generic training, while the “new training model” by [Acemoglu and Pischke \(1998, 1999\)](#) show that the incentive to train workers in generic training is restored under labor market imperfections. Despite these differences, the shared view is that training is beneficial to productivity, essentially by assumption.

A consistent view, supporting a positive effect of training on productivity, is shared by studies which examine training within the extended production function framework augmented to also include intangible (or knowledge) capital. Following the seminal contribution by [Corrado et al. \(2005\)](#), in this strand of research, training investment is part of a wider strategy to invest in intangibles. It contributes, in particular, to increase the component of intangible capital defined as “knowledge embedded in firm-specific human and structural resources”. Studies bringing the original macro-level framework to the firm-level, agree on a positive effect of intangibles on productivity (see [Webster and Jensen, 2006](#), for a review), in general, and about a notable effect of training, in particular ([Roth et al., 2023](#)).

A notable feature in the context of our study, common to all these “mainstream approaches to training”, is that the predicted positive effect of training on productivity is independent from firm size: micro, small, medium and larger firms are all equally expected to see their productivity increasing if investing in training. The empirical observation that, in reality, firms of relatively smaller size under-invest in training (or do not invest at all) has stimulated various explanations. A popular framework within the literature on SMEs – though not specifically targeted to micro firms – resorts to either “the ignorance explanation” or “the market forces explanation” ([Westhead and Storey, 1996](#)). The “ignorance explanation” highlights that smaller firms tend to consider training investment as far beyond the firms’ contingent needs and lack awareness of the potential returns, as they see training more as a cost than an opportunity to create value. The “market forces explanation” refers to market failures from either the supply or the demand side. The low market power and the significant turbulence usually featuring small firms, make training investment riskier and more uncertain in the outcome for this class of firms, compared to larger entities. The nature of the investment in training is perceived by smaller firms as involving long-term and delayed effects on performance, whereas these firms are usually more biased toward short-term returns than larger companies. Relatedly, micro-firms likely lack finance or prefer to devote the limited available resources to investment other than training. Still, while all

these factors may lead to sub-optimal or zero investment in training by firms of small size, training is not questioned to have a positive productivity pay-off: were policies able to remove such obstacles or misconceptions, restoring the correct incentives, then micro-firms which invest in training would surely reap the – by definition positive – returns of training on productivity.

A beneficial effect of training on productivity is also postulated in the “risk of poaching” explanation for low or zero provision of training by smaller firms. Directly deriving from the above-mentioned neoclassical theories of human capital accumulation in firms ([Mohrenweiser et al., 2019](#)), this explanation highlights the trade-off between improving productivity via investing in training, and the risk that trained employees may leave the firm shortly after training, particularly in the case of generic vs. firm-specific training. Despite firms of any size are exposed to poaching risk to a certain extent, firms of very small size may be particularly affected by that, since they typically lack the ability and the resources to offer attractive job, career and pay opportunities which larger firms may more easily afford, resulting into a preference for not investing in training in the first place.

Despite all these frameworks contribute to feed the common wisdom that training has a positive impact on firm-level productivity, a major challenge to this view is represented by approaches inspired to the resource-based/capability-based view of the firm ([Barney, 1991](#); [Dosi et al., 2000](#)) and incorporating the conceptualization of knowledge accumulation and innovation – broadly intended – as evolutionary processes ([Dosi, 2023](#)). In these approaches, it does not make sense to see human capital or knowledge/intangible capital as special inputs which, once augmented via a monetary investment in training, automatically contribute to productivity. Training is not intended as a means of accumulation of task-specific individual skills, but as related to the overall process of learning and innovation at the level of the firm as an organization. In fact, the locus of learning and knowledge accumulation is the firm as a collective organization, not the individual employee. The knowledge base of the firm is not the sum of the know-how or human capital of the individual employees (if anything like that kind of sum exists in these theories). It is rather a collective organizational property. In this respect, the distinction between generic and firm-specific training is meaningless: any training is essentially firm-specific, since the irreducible heterogeneity of knowledge across firms virtually never allows an employee trained in a firm to fit precisely in the same way in another workplace ([Koch and McGrath, 1996](#)).

As a result, the transmission mechanism going from enhancing employees’ knowledge and competences via training to productivity, is not guaranteed to be automatically positive. The outcome is heavily dependent from firm-specific characteristics, resting on idiosyncratic knowledge, routines and heuristics ([Dosi et al., 2021](#)). This reading of the role of training is also very much in tune with the view that comes from a complementary stream of the innovation literature, which conceives training as an investment that expands the firm’s absorptive capacity, easing the ability to acquire and exploit external knowledge. Differently from standard human capital theories, training operates here not as an intangible investment that may spur productivity, but rather as a key precondition factor, complementary to innovation, whose final effect on productivity is, once again, very much dependent on the specificity of the firm context.

As firm-specific characteristics and idiosyncratic processes of knowledge accumulation are central in non-mainstream, evolutionary approaches to innovation – broadly intended – in tune with the resource-based/capability-based theories, it is not surprising that it is within these frameworks that the specificity of micro-firms finds deeper recognition ([Kelliher and Reinl, 2009](#); [Sels et al., 2006](#)). In fact, micro-firms show peculiar characteristics which make them not just a mere scaling down of SMEs and larger firms. By their very nature, micro-firms are centered around a small group of core people or even a single person (the owner/manager). The core people shape firm’s decisions and performance in all the dimensions of firms’ activity ([Wennekers and](#)

Thurik, 1999; Liberman-Yaconi et al., 2010). In this context, the process of competence and knowledge accumulation through people becomes central, as the strength of micro-firms ultimately lies in the owner/manager and core-individuals' knowledge and capabilities (Rodrigues et al., 2022).

While such strong ties between people and firm outcomes could suggest that training is particularly attached to productivity performance in micro-firms, the very same specificity of these firms provides theoretical reasons why this may not be the case. First, the argument that training, as a device of knowledge absorption, is just a pre-condition for innovation and productivity, thus with no guarantee to deliver innovation and productivity outcomes, may be particularly relevant for micro-firms. In fact, their typically thin internal knowledge stock make them structurally more dependent on external knowledge (e.g., from suppliers, clients and local networks) than on formal internal R&D or technological innovations. Under this view, training in micro-firms may function more as a complementary investment that can only pass-through productivity when coupled with formal innovation investments, and less as a substitute for formal investment which directly spurs productivity.

Second, even sticking to the notion that improving core people's knowledge and capabilities may have a direct impact on productivity performance – beyond the indirect channel passing through formal technological innovation – training may not provide an obvious positive effect. On the one hand, training schemes may not serve the needs of smaller businesses, or even be detrimental, as they may result into too specialized staff, whereas multidimensional skills, creativity and flexibility are needed. As a result, knowledge accumulation and competence upgrading may be better coped through informal learning mechanisms such as on-the-job training, activity-based and cross-participated learning or coaching (Devins and Johnson, 2003; Kotey and Slade, 2005; Lopez et al., 2019). Moreover, as suggested by managerial theories that are converging toward a “bundle view” of the effect of HRM on small firm performance (Jiang et al., 2013; Rodrigues et al., 2022), the contribution of training to productivity may be hampered in micro-firms, if not appropriately integrated into the complexities of the whole HRM strategy — whereof training provision can be seen as a component. In fact, whether formalized or not, HRM practices are not necessarily less sophisticated in small firms than in larger entities (Sels et al., 2006).

2.2. Training and productivity in micro-firms: empirical studies

Notwithstanding the theoretical argument that micro-businesses represent a peculiar category of firms, empirical studies are essentially silent on the training–productivity link in micro-firms. Well-known and highly cited works which contribute to the common wisdom that training fosters productivity (Bartel, 1994; Black and Lynch, 1996; Dearden et al., 2006; Konings and Vanormelingen, 2015), either do not cover micro-firms at all or do not separate them from the wider categories of small firms and SMEs. In this studies, the estimated productivity gains range in between 2 and 15 percent, but comparability across works is difficult due to wide heterogeneity in the empirical approaches (see Martins, 2021, for a review). There are indeed wide differences in terms of context (countries and sectors), estimation methods (simple panel estimates vs. methods addressing endogeneity, such as instrumental variables or quasi-experimental estimation) and measures of training (expenditures vs. percentage of trained employees, specific vs. general, private vs. public grants).

Empirical studies of the Italian case do not represent an exception to the general rule that micro-firms are not analyzed separately. Mancinelli and Mazzanti (2009) document that training expenditures per employee have a positive but very small (about 1% increase) effect on productivity in a sample of SMEs active in the Emilia-Romagna region. Dosi et al. (2021) document a similarly weak effect of training on productivity in a large sample representative of the population of

firms with less than 50 employees, but they do not provide separate estimates for the subset of micro-firms.

The few empirical studies which dedicate specific attention to micro-firms, separating them from other firms, do not help to advance our understanding of the training–productivity link in micro-firms. On the one hand, there are studies which take productivity as the outcome performance variable of interest, but do not isolate the effect of training from other components of HRM in micro-firms, consistently with the theoretical shift in the managerial micro-business literature toward the “bundle view” of HRM (see the extensive review in Jiang et al., 2013). On the other hand, there are the few studies which do try to isolate the effect of training – or, more broadly, of knowledge accumulation – on micro-firms' performance, but do not take productivity as the performance outcome of interest. Westhead and Storey (1996) and Storey (2004) provide reviews of early studies, mostly based on data about governmental programs in the UK context, concluding that the link between training and performance (survival and growth dynamics, in particular) is not well established. This conclusion is in line with the unstable and weak relation between training and various growth variables – among which productivity *growth* but not productivity *levels* – documented in Cosh et al. (1998). Coad and Tamvada (2012) provide some indirect evidence in the context of an emerging economy (India), showing that different forms of sourcing external technical knowledge may be important for growth dynamics of small-micro-firms. Henley and Song (2020) study micro-firms, again in the UK context, and document that knowledge upgrading in various forms may boost productivity of micro-businesses at least indirectly, by primarily exerting a positive impact on their participation to international markets, which in turn fosters productivity.

2.3. Research question and working hypotheses

Overall, from the existing literature, the received wisdom that training should foster productivity of micro-firms, as it does in firms of larger size, remains theoretically unclear and empirically untested.

The empirical estimation of the effect of training on productivity in micro-firms, could entail two possible strategies. One path could be to select data covering both micro and larger firms, and then compare the productivity performance of micro-firms which train vis-à-vis productivity outcomes of larger firms which also train. In this paper, we tackle the complementary, and more basic research question. We examine whether there exists a productivity premium for micro-firms that do invest in training vis-à-vis micro-firms that do not.

As the existing theoretical and empirical literature is inconclusive, what we could expect to find is not straightforward to predict a-priori. On the one hand, one could simply resort to the generic, somewhat simplistic intuition that training always boosts productivity, no matter firm size. Thus, micro-firms which train should be more productive than non-training micro-firms, just as any firm investing in training is expected to be more productive than in the counterfactual situation of not investing. In fact, this positive relation may be particularly strong in micro-firms, considering the people-centered nature of these enterprises, implying a particularly strong dependence of business performance on the knowledge and competences of the owner/manager and core employees.

On the other hand, however, the same peculiar characteristics of the micro-firms may lend support to the opposite prediction that the “control group” of micro-firms that do not spend in training, may turn out as equally productive as or even more productive than firms that invest in training. For a number of reasons. A first reason is suggested by the “risk of poaching” argument. Even if training is assumed to foster productivity in this framework, the reality may be that micro-firms which train see their trained employees constantly leaving the firm, and are thus left without any productivity return from their investment in training. Second, as suggested by the resource-based/capability-based view, in the context of micro-firms it is reasonable to hypothesize that

other modes of competence upgrading/updating, both at the individual level and at the level of the organization, may prove as particularly effective, delivering a larger productivity pay-off than training. Third, the productivity losses for micro-firms which invest in training, may be larger than the productivity gains connected to training. In fact, investing in training comes obviously with costs. These costs are not only monetary, but also involve time and coordination efforts distracted from other activities and devoted to identifying and/or designing valuable training programs, and managing and attending them. The “ignorance” and “market forces” explanations convey the notion that micro-firms weigh these costs too much while failing to recognize the potential gains, or perceive them as unrecoverable, and thus they do not invest, due to strong market failures. But what if – reversing the usual narrative – training is merely a waste of resources and not suited to improve productivity in the specific context of micro-firms? Were this reasoning to have an empirical traction, we would observe a zero or even negative productivity premium for micro-firms that spend in training vis-a-vis micro-firms that do not. Our baseline analysis, in Section 4, tackles whether training has an overall positive effect or not in the data. Then, in the extended analysis of Section 5, we explore the heterogeneity of the effects by firm size, intensity of firms’ investment in tangible and intangible assets, and sector of activity, in turn shedding light on the complementarities between training and these dimensions of heterogeneity.

3. Data and descriptive evidence

The empirical analysis of this study exploits the access to the TREND database. This is an administrative dataset originated from the collaboration between the Italian national statistical office (ISTAT) and one of the main associations of SMEs in Italy (Confederazione Nazionale dell’Artigianato e Piccola-Media impresa, CNA).² The TREND project started as a pilot in 2006 in the Marche region with the constitution of an observatory on small and micro enterprises. It was later extended to three other regions of Central Italy (Emilia-Romagna, Umbria and Toscana), with the purpose to gather firm-level information about key performance and structural characteristics of CNA associates, thereby integrating the information regularly collected by ISTAT in the official statistics, which are less detailed for micro-firms below 10 employees than for larger entities. The TREND database is constructed to be representative of the population of firms at the regional level, by firm size and sector of activity, through application of appropriate weights to align the data coming from CNA associates with the official register of the population of active firms (ASIA-Archivio Statistico Imprese Attive) maintained by ISTAT.³

The TREND dataset made available to us covers 21,044 micro-firms active in Marche, Toscana and Umbria over the period 2013–2017.⁴ The included variables encompass the following firm attributes, for each firm-year pair: number of employees (as the average number of employees over the calendar year, including the owner/manager), macro-sector of activity (at 1-digit NACE level), geographical region of

Table 1

Descriptive statistics — estimation sample.

(A) TREND variables:			
	N.Obs	Mean	Std. Dev.
Train (Training dummy)	42,658	0.10	0.30
– Training expenses (if Train = 1)	4,329	572.99	1,181.46
Size (N.employees)	42,658	1.93	1.50
Total sales	42,658	85,098.25	168,291.10
Export revenues	42,658	1,111.36	26,986.03
Tangible investment	42,658	5,975.29	63,044.72
Intangible investment	42,658	327.19	14,194.97
(B) Constructed variables:			
	N.Obs	Mean	Std. Dev.
Productivity	42,658	4,1705.95	64,278.38
– Productivity (Train = 0)	38,329	40,239.55	64,737.25
– Productivity (Train = 1)	4,329	54,689.56	58,487.00
Sales growth	42,658	0.14	4.46
Export dummy	42,658	0.02	0.15
Tangibles intensity	42,658	2,357.22	24,517.84
– Tangibles dummy	42,658	0.30	0.46
Intangibles intensity	42,658	114.64	6,090.68
– Intangibles dummy	42,658	0.02	0.15

Notes: The table reports descriptive statistics on the estimation sample. Train is a dummy for having non-zero training expenses; Productivity is defined as sales (revenues) per employee; Sales growth is the growth rate of sales (revenues) between $t-1$ and t ; Tangible and Intangible intensity are defined as tangible and intangible investment divided by number of employees, respectively.

location, total sales (revenues), the value of export, training expenses and investments in tangible and intangible assets.⁵ Although the list may seem rather limited, compared to the wider set of balance-sheets items usually available for larger firms from official statistics or commercial datasets, the TREND data are actually quite rich considering that we are dealing with micro-firms. In fact, even the population-wise censuses maintained by statistical offices around the world, either do not include micro-firms at all or, when they do, record information only about very few, basic firm characteristics. Specifically, the TREND dataset improves upon the ASIA register maintained by ISTAT, as the latter just records employees, sales, sector and geographical location.

The variable which is central to our purposes is the monetary value of training expenses incurred by each firm per year. This represents a rather unique opportunity for research, since information about training is not usually available in the typical datasets on micro-firms. Specifically, the training expenses recorded in TREND refer to the costs incurred to update employees’ knowledge and competencies related to their function and tasks in the firms. They thus capture investment in specific training, while excluding expenses undertaken to finance generic training or costs related to mandatory training (in compliance with regulations of various kind, such as health&safety-related training).

In our main analysis, we exploit the information on training expenses to define the treatment status, as a dummy variable (*Train*) distinguishing micro-firms with positive vs. zero training investment. The information on annual training expenses (continuous variable) reported in the raw data is instead exploited in the robustness analysis.

As the dependent variable, we compute a simple labor productivity index (*Productivity*), defined as sales (revenues) per employee, while the other firm-level attributes provided in the TREND data are variously used as auxiliary variables in the identification of the training effect on productivity. In particular, starting from the available variables, we compute the following (for each firm and year): sales growth (defined as the growth rate of sales over two consecutive years), to proxy for generic market performance; tangible and intangible intensity (respectively defined as investment in tangible and intangible assets scaled by

² The data were accessed thanks to a research agreement between the authors, Scuola Superiore Sant’Anna, ISTAT and local CNA offices, under a standard mandatory condition of censorship of any individual information.

³ Documentation about the TREND dataset is unfortunately available only in Italian. Examples of non-technical report exploiting the TREND data can be found on-line in the websites of regional units of CNA (e.g. at <http://www.trendmarche.it>), while a more technical discussion of data construction and representativeness is provided in Palmieri (2016). Additional documentation on development and characteristics of the TREND database was accessed by the authors within the framework of their joint research project with ISTAT and CNA, and can be provided upon request.

⁴ As the TREND project was still under development at the time we submitted the data access request, the data about the fourth region covered by the project (Emilia-Romagna) were not available.

⁵ Nominal variables are all deflated at the origin by appropriate sectoral deflators.

number of employees), which we use in various model specifications either directly in levels or as dummies for positive vs. zero tangible or intangible investments; an export dummy, recording whether the firm is selling abroad or otherwise.

Considering that one year of data (2013) is lost to compute sales growth and that some firms may become larger than 10 employees over the sample time window (i.e., they cease to be micro-firms), while some missing firm-year observation may affect some of the variables included in the empirical analysis (see next section for details on regression models), the final working sample available for estimation includes 42,658 firm-year observations referring to 16,797 firms over the period 2014–2017.

Table 1 reports descriptive statistics computed over the estimation sample. Panel A refers to the set of variables originally recorded in TREND, while in Panel B we report about the variables that we constructed from the data. From Panel A, we observe that only about 10% of the sample records non-zero training expenditures (mean of the Train dummy is 0.10). This figure is in line with the stylized fact that training is rare among micro-firms. It also supports our choice to exploit information about training expenses in a dummy form, since crucial variation is not in the intensity of training, but in whether firms do invest or not in training in the first place. The statistics about number of employees are also interesting. Considering that, in the TREND data, the recorded employees include the owner, the observed average number of employees close to 2 (cf. average size ≈ 1.93) and the median of 1, imply that the sample is essentially made of firms centered around the owner/manager and a very small group of core employees. Indeed, by disaggregating further, we find that about 86% of the firm-year observations have up to three employees and about 95% has up to 5 employees, while just about 2% have more than 7 employees. From Panel B, we see that, despite positive training is relatively rare, training firms (Train = 1) are on average more productive. This is a preliminary indication that training may yield a productivity premium, yet formal regression analysis is required before drawing solid conclusions.

4. Main analysis and robustness

As an initial framework to assess the impact of training on productivity, we start from the following baseline regression model

$$\ln(\text{Productivity})_{i,t} = \beta_0 + \beta_1 \text{Train}_{i,t} + \beta_2 X_{i,t} + \eta_t + \eta_s + \eta_r + \epsilon_{i,t}, \quad (1)$$

where, for each firm i and year t , *Productivity* stands for labor productivity (sales per employee), *Train* is the dummy for positive training expenditures, while the set of control variables $X_{i,t}$ includes all the firm-level variables mentioned before: sales growth, firm size in terms of number of employees (in logs), tangible and intangible intensity (in logs), and the dummy for exporting firms. A full set of year (η_t), sector (η_s) and region (η_r) fixed-effects is also included.

Table 2 shows simple OLS and firm fixed-effects (FE) estimates of Eq. (1). The results support a positive and significant productivity premium for firms that spend in training, over firms that do not. When we add firm fixed-effects, delivering more credible estimates, the estimated productivity-premium reduces considerably but it still remains statistically significant and sizable.

Since most of the unobserved and potentially relevant firm-specific attributes –such as managerial practices or attitude, owner's or core people's characteristics– arguably remain constant over the short estimation time-span of 4 years, firm fixed-effects are absorbing the large part of the related potential bias. Still, the estimates may suffer from standard residual endogeneity concerns, regarding endogenous selection into the training group.

In order to address these concerns, and improve identification, we rely on the IPWRA method, which allows to circumvent the lack of exogenous variation in our data (for details, see Wooldridge, 2007; Słoczyński et al., 2022; Cappelletti and Giuffrida, 2022). The logic of the methodology goes essentially in two steps. In the first step, the

Table 2
Simple OLS and firm FE estimates.

	OLS	Firm-FE
Training dummy	0.534*** (0.019)	0.179*** (0.018)
Sales growth	0.021*** (0.007)	0.023** (0.010)
ln(Size)	0.039** (0.017)	−0.569*** (0.051)
ln(Tangible intensity)	0.102*** (0.002)	0.025*** (0.002)
ln(Intangible intensity)	0.001 (0.010)	0.007 (0.008)
Export dummy	0.707*** (0.046)	0.315*** (0.071)
Obs.	42,658	42,658

Notes: OLS and firm fixed-effects (Firm-FE) estimates of regression model in Eq. (1). OLS estimates include year, region and sector fixed-effects, while region and sector fixed-effects are excluded from firm FE estimates as they are time-invariant. Cluster-robust standard errors (at firm-level) in parenthesis. Asterisks denote significance levels (* at 10%, ** at 5%, and *** at 1%).

predicted probability (p-score) of being treated based on observable characteristics, is obtained for each firm. In the second step, the (square root of the) inverse of the p-scores obtained from the first step is used as a weight in two adjusted regression separately estimated across treated and untreated units, where the dependent variable is the outcome of interest. Then, an estimate of the ATE is provided by taking the difference between the averages of potential outcomes obtained from the second-step regressions for re-weighted treated vs. re-weighted untreated units. On the one hand, the regression adjustment (RA) accounts for different trends between treated and untreated observations by allowing for different coefficients for the two groups. On the other hand, the inverse probability weighting (IPW) achieves comparability across treated and untreated units, balancing out differences in observable characteristics by giving high weight to treated observations having a relatively low probability of treatment, and to untreated units having a high probability of treatment. Essentially, this creates an artificial counterfactual sample where untreated units are equally likely to be in the treated group, by downsizing two types of observations: (i) observations which are treated and have high predicted probability of treatment, and (ii) observations which are not treated and have low predicted probability of treatment. The key property and advantage of the IPWRA over related methods is that the IPWRA is doubly-robust: while IPW estimates are valid only under correct specification of the propensity scores (treatment) equation and RA estimates are valid only under correct specification of the relation between the outcome and the propensity scores, the IPWRA allows for mis-specification of either the outcome or the treatment equation.

In our context, the final outcome is productivity and the treatment corresponds to belonging to the group of micro-firms which train. We specify a first-step Probit equation of the form

$$\Pr(\text{Train}_{i,t} = 1) = \Phi(\alpha_0 + \alpha_1 Z_{i,t}), \quad (2)$$

and use the predicted p-scores ($\hat{p}_{i,t}$) to generate the weights entering the following second-step adjusted regressions

$$\ln \text{Productivity}_{i,t}^1 = \gamma_0^1 + \gamma_1^1 X_{i,t} + \eta_t + \eta_s + \eta_r + u_{i,t}^1$$

with weights $\sqrt{\frac{1}{\hat{p}_{i,t}}}$ (3)

$$\ln \text{Productivity}_{i,t}^0 = \gamma_0^0 + \gamma_1^0 X_{i,t} + \eta_t + \eta_s + \eta_r + u_{i,t}^0$$

with weights $\sqrt{\frac{1}{1 - \hat{p}_{i,t}}}$ (4)

which mimic Eq. (1) above (same covariates and fixed-effects) but are estimated respectively on re-weighted treated and re-weighted

Table 3

Baseline IPWRA estimates, first-step results and covariate-balance.

Observations	Raw	Weighted			
Total Obs.	42,658	42,658			
– Treated Obs.	4,329	21,308.8			
– Control Obs.	38,329	21,349.2			
Variable	Probit Coefficients	Raw Std diff	Weighted Std diff	Raw var ratio	Weighted var ratio
Sales growth	0.0001 (0.001)	0.021	−0.002	0.960	0.646
Single-empl. dummy	−0.711*** (0.022)	−0.770	−0.003	0.720	1.000
Tangible dummy	0.314*** (0.019)	0.435	0.006	1.237	1.005
Intangible dummy	0.118** (0.049)	0.147	0.001	2.247	1.005
Export dummy	0.033 (0.062)	0.070	−0.007	1.518	0.955
Covariate-balance test: Test stat = 3.270 p-value = 0.7742					

Notes: The table reports number of observations in the raw and re-weighted sample, together with first-step Probit coefficients and covariate-balance diagnostics (Standardized differences and Variance Ratios). Cluster-robust (by firm) standard errors in parenthesis, asterisks denote significance levels (* at 10%, ** at 5% and *** at 1%). Covariate-balance test is the Imai and Ratkovic (2014) over-identifying restriction test for covariate balance: χ^2 test statistic and corresponding *p*-value are reported.

untreated observations (indicated with the 1 and 0 superscripts, respectively). The ATE of training is then obtained as the difference between the average fitted outcomes from the two second-step regressions.

The crucial requirement for the IPWRA estimates to be credible is that the *Z* covariates included in the first-step Probit achieve covariate balance, which is a necessary condition to meet the common-support assumption underlying the procedure. Although including a reasonably comprehensive and sensible set of regressors in the Probit is a natural way to proceed, there is flexibility in designing the Probit specification, provided that the included variables achieve covariate balance. After experimenting with different specifications, starting from the relatively small set of variables available from TREND data, we include in *Z* the following: sales growth, two dummies for zero vs. positive tangible and intangible investments, a dummy for being a single-employee enterprise and the export dummy.

Table 3 reports the first-step Probit coefficients and the covariate-balance analysis. As conventional in the literature, we show standardized differences in covariates and variance ratios of covariates across treated and control observations, in the raw data vs. after re-weighting. We obtain quite satisfactory results. Weighted standardized differences are all quite below (in absolute value) 0.25, which is the benchmark upper bound for covariate balance, while the variance ratios after weighting are all (but for sales growth) equal or very close to the benchmark value of 1 (Imbens and Rubin, 2015; McCaffrey et al., 2004; Stuart and Yim, 2010). We also report the results of the Imai and Ratkovic (2014) over-identifying restrictions test for covariate balance: the null of covariate balance cannot be rejected, providing strong support to the credibility of the estimates.

Table 4 displays OLS (OLS-IPWRA) and firm fixed-effects (FE-IPWRA) estimates of the second-step weighted regressions, together with the corresponding ATEs of training. While FE analysis is non-standard in applications of the IPWRA methodology, it is nevertheless worth attempting, as an additional way to absorb firm-specific time-invariant heterogeneity in the second-step regressions. In fact, while the structural part of endogenous selection into treatment status is absorbed by the *p*-score weighting, absorbing unobserved time-invariant heterogeneity is important, as the IPWRA relies upon a standard CIA assumption and the TREND data only allow for a limited set of variables to include in the estimates. By adding firm fixed-effects, our identification exploit that most of the potentially relevant omitted

Table 4

Baseline IPWRA estimates, second-step results.

	OLS-IPWRA		FE-IPWRA	
	(1) ln Prod ⁰	(2) ln Prod ¹	(3) ln Prod ⁰	(4) ln Prod ¹
Sales growth	0.020*** (0.007)	0.011** (0.005)	0.023 (0.015)	0.021* (0.012)
ln(Size)	0.050** (0.019)	−0.017 (0.027)	−0.302*** (0.055)	−0.805*** (0.062)
ln(Tangible intensity)	0.108*** (0.003)	0.058*** (0.004)	0.030*** (0.002)	0.018*** (0.003)
ln(Intangible intensity)	−0.003 (0.012)	0.037*** (0.013)	−0.003 (0.009)	−0.012 (0.011)
Export dummy	0.755*** (0.049)	0.402*** (0.103)	0.418*** (0.080)	0.100** (0.044)
Obs.	38,329	4,329	38,329	4,329
ATE	0.614*** (0.026)		0.177*** (0.012)	

Notes: The table reports OLS and firm fixed-effects (FE) estimates of the IPWRA second-step weighted regressions, and the corresponding ATEs of training. OLS regressions also include year, region and sector fixed-effects, while region and sector fixed-effects are excluded from the firm-FE models as they are time-invariant. Robust standard errors in parenthesis clustered by firm. For the OLS estimates, standard errors are corrected to account for the two-step procedure via the STATA *teffects* command. As the correction is not available for FE estimation, corrected standard errors clustered by firms are obtained via bootstrap (200 replications). Asterisks denote significance levels (* at 10%, ** at 5% and *** at 1%).

factors (such as managerial ability or strategy or owner's /core people's characteristics) are indeed arguably constant over the short estimation time-period.⁶

The ATE of training obtained from both the OLS-IPWRA and the FE-IPWRA estimates is positive and statistically significant. As one might expect, the FE-IPWRA estimate is smaller. In this more credible set-up, absorbing time-invariant heterogeneity, the point estimate implies a difference in productivity of about 19.4% (0.18 log-points), on average, between micro-firms that invest in training vis-a-vis the counterfactual case of micro-firms that do not invest in training.

In order to corroborate the robustness of the main results, we complement the baseline IPWRA analysis with two additional exercises.⁷

A first issue concerns the decision to condense the information about training expenditures into a 0–1 indicator of training status. This choice is motivated primarily by the observation that only about 10% of the sample has non-zero training expenditures (recall Table 1). This feature of the data implies that the bulk of variation is not in how much firms invest in training, but in whether they do that or not, in turn supporting that a more reliable statistical signal could come from distinguishing who is investing vs. who is not investing in training. For a matter of robustness and also for comparison with the practice in the empirical literature on training effects (in larger firms, not in

⁶ Regarding the OLS-IPWRA analysis, the coefficient estimates and robust standard errors clustered by firm, are obtained via the STATA command *teffects*, which appropriately corrects the standard errors (on the regressors and on the ATE) to account for the two-step procedure. Regarding the FE-IPWRA analysis, the procedure was manually implemented. After computing the *p*-scores from the Probit selection Eq. (2), the second-step involves a weighted OLS performed over a usual FE-within group (within-firm) transformation of the outcome Eqs. (3) and (4). In this case, robust standard errors are computed via bootstrap, with 200 runs, which were enough to achieve convergence.

⁷ In a further, unreported, robustness check, we estimated a different second-step model of the IPWRA, where we controlled for size via three dummies for 1–2, 3–5 and 6+ employees, in place of the number of employees. The estimated ATE of training was in line with the baseline analysis. The results are available upon request. Also, the magnitude and significance of the baseline IPWRA results are confirmed through additional sensitivity analysis involving the application of alternative estimators for binary treatment, reported in Appendix A.

Table 5
Models with training expenditures.

	Training expenses			Training intensity		
	OLS (1)	Firm-FE (2)	GMM-SYS (3)	OLS (4)	Firm-FE (5)	GMM-SYS (6)
ln(Train. expenses)	0.092*** (0.003)	0.032*** (0.003)	0.041*** (0.013)			
Train. intensity	–	–	–	0.006*** (0.0001)	0.002*** (0.00004)	0.010** (0.005)
Sales growth	0.021*** (0.007)	0.023** (0.010)	0.076 (0.059)	0.021*** (0.007)	0.023** (0.010)	0.049 (0.063)
ln(Size)	0.038** (0.017)	–0.570*** (0.051)	–1.005*** (0.262)	0.098*** (0.017)	–0.557*** (0.051)	–0.882*** (0.265)
ln(Tang. intensity)	0.102*** (0.002)	0.025*** (0.002)	0.537*** (0.069)	0.105*** (0.002)	0.026*** (0.002)	0.568*** (0.088)
ln(Intang. intensity)	0.001 (0.001)	0.007 (0.008)	–0.309*** (0.054)	–0.0001 (0.010)	0.007 (0.008)	–0.304*** (0.105)
Export dummy	0.706*** (0.046)	0.315*** (0.071)	0.421*** (0.092)	0.706*** (0.045)	0.320*** (0.071)	0.334*** (0.117)
Obs.	42,658	42,658	42,658	42,658	42,658	42,658
GMM-SYS diagnostics:	AR(1) test: AR(2) test: Hansen's J:	Test stat. = –7.65 Test stat. = –0.88 Test stat. = 16.75	p-value = 0.0001 p-value = 0.378 p-value = 0.334	AR(1) test: AR(2) test: Hansen's J:	Test stat. = –4.33 Test stat. = –1.14 Test stat. = 12.15	p-value = 0.0001 p-value = 0.256 p-value = 0.353

Notes: The table reports OLS, firm FE (Firm-FE) and two-step GMM-SYSTEM (GMM-SYS) estimates of Eq. (1) taking either training expenses (in logs) or training intensity (expenses per employee) in place of the training dummy. OLS estimates include year, region and sector fixed-effects, while region and sector fixed-effects are excluded from firm FE and GMM-SYS estimates as they are time-invariant. Robust standard errors in parenthesis, clustered by firm. Asterisks denote significance levels (* at 10%, ** at 5% and *** at 1%). We also report results (test-statistics and associated p-values) of the Arellano–Bond tests for serial correlation of order 1 and order 2 (AR(1) and AR(2)), and the robust Hansen's J test of overidentifying restrictions for instruments validity.

micro-firms), it is nonetheless interesting to also exploit the information on training expenditures (continuous variable) as recorded in the raw data. We examine two alternative specifications of the initial model in Eq. (1), respectively including training expenses or training intensity (expenses per employee) in place of the training dummy *Train*. Since the IPWRA methodology does not accomplish for continuous treatment, we tackle potential endogeneity of training through a standard (two-step robust) GMM-SYSTEM estimator. The results are reported in Table 5, together with simple OLS and firm fixed-effects estimates of the corresponding models, for benchmarking of the GMM-SYSTEM estimates.⁸ The estimates confirm a positive and significant productivity premium of training. This is also reassuring from a statistical viewpoint, since the GMM-SYSTEM estimator exploits an IV-like identification strategy, complementary to the main IPWRA approach to endogenous selection into treatment.

The further variation of the baseline analysis that we explore, deals with the possible presence of lagged effects of training. This additional analysis is informative under two respects. On the one hand, one could argue that it may take some time for the full effect of training to manifest, due to the time required before new knowledge and competencies of trained people are "absorbed" into firms' organization and practices, eventually yielding improved efficiency. On the other hand, one could be interested in uncovering the degree of persistence of the productivity effect of training over time. For instance, productivity premia may reduce over time if training-related knowledge becomes rapidly obsolete or if – in line with the poaching risk hypothesis – trained employees leave shortly after training. To explore these dynamics, we modify our baseline IPWRA analysis, looking now at the effect of 1-year lagged training status. In practice, we obtain lagged propensity scores p_{t-1} from a Probit of lagged training status $Train_{t-1}$ against lagged predictors Z_{t-1} , and then use the lagged propensity scores to compute the corresponding weights entering the second-step outcome equations of the IPWRA procedure. The estimated ATEs,

⁸ The instruments set used in the GMM-SYSTEM includes various lags of size and intangibles. Validity is supported by the standard Arellano–Bond tests for serial correlation of order 1 and 2 (labeled as AR(1) and AR(2) in the Table) and the Hansen's J test of overidentifying restrictions for instruments validity.

Table 6
1-year lagged training dummy, IPWRA second-step results.

	OLS-IPWRA		FE-IPWRA	
	(1) ln Prod ⁰	(2) ln Prod ¹	(3) ln Prod ⁰	(4) ln Prod ¹
Sales growth	0.046*** (0.014)	0.247*** (0.092)	0.010 (0.013)	–0.013 (0.033)
ln(Size)	0.059*** (0.021)	0.033 (0.044)	–0.386*** (0.055)	–0.573*** (0.104)
ln(Tangible intensity)	0.106*** (0.003)	0.074*** (0.007)	0.036*** (0.003)	0.032*** (0.007)
ln(Intangible intensity)	0.038*** (0.011)	0.039*** (0.014)	0.015 (0.011)	0.007 (0.011)
Export dummy	0.738*** (0.053)	0.354*** (0.120)	0.372*** (0.083)	0.113 (0.094)
Obs.	22,854	2,714	22,854	2,714
ATE	0.422*** (0.039)		0.088*** (0.021)	

Notes: The table reports OLS and firm fixed-effects (FE) estimates of the IPWRA second-step weighted regressions, and the corresponding ATE of training. OLS regressions also include year, region and sector fixed-effects, while region and sector fixed-effects are excluded from firm FE models as they are time-invariant. Robust standard errors in parenthesis, clustered by firm. For the OLS estimates, standard errors are corrected to account for the two-step procedure via the STATA *teffects* command. As the correction is not available for estimation with firm-FE, corrected standard errors clustered by firm are obtained via bootstrap (200 replications). Asterisks denote significance levels (* at 10%, ** at 5% and *** at 1%).

reported in Table 6, are lower than in the baseline analysis, but still positive and statistically significant, thus confirming the beneficial role of training in boosting productivity of micro-firms.⁹ From a statistical viewpoint, this is reassuring because estimates with lagged training status reduce reverse causality concerns related to simultaneity in the baseline analysis. From a more substantive standpoint, the observed reduction in the productivity premium suggests that constant training might be necessary to keep the productivity-enhancing effect of training

⁹ The related first-step Probit results and covariate-balance analysis are reported in Appendix B, see Table 13.

constant over time, as training become obsolete or new employees join the firm while some are poached from outside. In any case, whatever the explanation, the analysis of lagged effects adds to the main analysis that the productivity advantage of training does not completely vanish over the very short-run.

5. Heterogeneity and complementarity analysis

The joint reading of the findings presented in Section 4 establishes confidence that training has a significant, positive effect on productivity of micro-firms. This effect is sizable, around 19 percentage-points, according to the more credible FE-IPWRA results.

We now move to explore whether the productivity return to training varies along four dimensions which our data allow to consider: firm size, intensity of intangible and tangible investments, and sectoral membership to manufacturing. This analysis allows us to address whether training firms see their productivity premium reinforced or reduced along these heterogeneity dimensions, thus providing at least indirect evidence on some channels, attached to these dimensions, which may underlie the baseline effect of training.

In a standard regression framework connecting outcomes to a treatment dummy (training status in our case), heterogeneity or channels analysis is customarily investigated through interaction effects between treatment status and the dimensions of interest. This is not viable within the IPWRA framework. The training dummy is used in the first-step Probit to compute p-scores and obtain associated weights, but it does not directly enter in the second step adjusted regression. It cannot be interacted with any other variable.

To examine heterogeneity within our IPWRA framework, we move from the baseline IPWRA single-treatment setting to a multivalued setting (Cattaneo, 2010; Wooldridge, 2010; Cattaneo et al., 2013), in which different groups of firms are identified, by interacting the training dummy with dummies that split the sample according to each of the four above-mentioned dimensions: a dummy for single-employee (vs. multi-employee) firms; a dummy for positive (vs. zero) intangible investment; a dummy for positive (vs. zero) tangible investment; a dummy for manufacturing (vs. non-manufacturing) firms. Accordingly, for each dimension of interest, we define four mutually exclusive groups that identify: (i) “baseline” firms, which do not train and do not display the focal dimension considered (e.g., they do not invest in tangibles); (ii) “dimension only” firms, which do not train while they exhibit the focal dimension (e.g., do not train while they invest in tangibles); (iii) “training only” firms, which train while they do not display the focal attribute (e.g., train and do not invest in tangibles); (iv) “combined group” firms, which both train and also display the focal characteristic (e.g., train and invest in tangibles).

Then, for each training $\times$ attribute pairing, we modify the IPWRA procedure, in particular the more solid estimator with firm fixed-effect, as follows. In the first step, we specify a Multinomial Logit model –the standard approach for estimating generalized propensity scores in a multivalued framework (Imbens, 2000; Cattaneo, 2010)– to estimate each firm’s probability of belonging to each of the four groups as a function of the observed covariates. This yields estimated propensity scores, which are then used to construct the weights for the second stage outcome regressions with firm fixed-effects, thus correcting for non-random selection into each group, in the same spirit of the IPWRA procedure applied in Section 4.

In the second step, a separate outcome equations is estimated for each of the four groups (“baseline”, “dimension only”, “training only” and “combined”), with each observation weighted by the corresponding inverse propensity weights. This yields four group-specific fitted models, which are then used to predict potential productivity outcomes under each of the four “regimes”. By averaging these predictions, one obtains doubly-robust estimates of the ATE of each group, relative to an omitted baseline reference group, which for us will be the “baseline” category as defined above. As in Section 4, since we need to account

for the fact that propensity scores feeding into the second-step weights are themselves estimated, corrected standard errors on the estimated ATEs (as well as the associated significance levels) are obtained via a firm-level clustered bootstrap procedure (200 replications).

A key advantage of the multivalued-grouping structure is that it allows to formally test for complementarity (supermodularity) or substitutability (submodularity) between training and each dimension of interest, an approach in the same spirit as the broader literature on combined treatment (Linden et al., 2016; Cattaneo, 2018; Ao et al., 2021). Denoting with $ATE(combined)$, $ATE(train\ only)$, and $ATE(dimension\ only)$ the estimated ATEs of each group relative to the reference group of “baseline firms”, we construct the contrast:

$$ATE(compl) = ATE(combined) - ATE(train\ only) - ATE(dimension\ only) \quad (5)$$

Under **additivity**, the effect of combining training with the focal dimension, labeled as $ATE(combined)$, would equal the sum of the individual effects $ATE(train\ only) + ATE(dimension\ only)$, so that $ATE(compl)$ would be statistically equal to zero. That would signal that training has no interaction with the focal heterogeneity dimension, hinting that the productivity effect of training is not affected by that dimension. A positive and significant $ATE(compl)$ would indicate **complementarity** between training status and the focal dimension of interest: the effect of combining them yields more productivity than the sum of their separate effects. Conversely, a negative and significant $ATE(compl)$ would signal **substitutability**: training “competes” with the focal dimension of interest for the same underlying productivity margin, so that their combined effect on productivity is less than the sum of their individual effects.

Because $ATE(compl)$ is itself a linear combination of the bootstrapped group-specific ATEs, the associated standard errors, confidence intervals and significance levels are obtained directly from the firm-clustered bootstrap distribution. This ensure that inference on the complementarity test properly accounts for the estimation uncertainty in both the first-step estimates of propensity scores and the second-step outcome regressions.¹⁰

5.1. Training and firm size: single- vs. multi-employee firms

Regarding heterogeneity of training effects across micro-firms of different size, we distinguish distinction between single-employee vs. multi-employee firms. This is particularly meaningful in our context. In fact, single-employee firms are entities where the owner/manager is the only person in the firm, while multi-employee firms are more “complex” organizations where additional core people are present.¹¹ If this increased complexity shaped how training translates into productivity – for instance, through delegation and diffusion of knowledge and competencies across a team rather than concentration in a single person

¹⁰ In an earlier version of this work, we addressed heterogeneity analysis through a split-sample analysis, applying the IPWRA procedure of Section 4 to obtaining separate ATEs of training within sub-samples defined by each dimension (e.g., single-vs. multi-employee firms). While the main results were broadly in line with the multi-valued approach presented here, that approach bears a limitation: comparing two separately estimated ATEs, each from a different sub-sample, does not deliver a formal statistical test of whether training and each focal dimension interact. It can only suggest, informally, whether point estimates look similar or different across sub-samples. The multivalued framework allows for a formal statistical test of complementarity, with inference drawn from a single bootstrap distribution that reflects the joint estimation uncertainty of training and focal-dimension status, simultaneously. Results of the earlier split-sample approach are available from the authors upon request.

¹¹ TREND data comply with this distinction: firms recorded as having just one employee are entities with just the owner/manager working there.

Table 7

Training and firm size — Multi-valued FE-IPWRA results.

Panel A — Second-step results:				
	Single-empl., no training	Single-empl., training	Multi-empl., no training	Multi-empl., training
Sales growth	0.081*** (0.023)	0.050 (0.066)	0.015 (0.020)	0.013 (0.015)
Tangible intensity	0.031*** (0.003)	0.026*** (0.006)	0.028*** (0.003)	0.012*** (0.003)
Intangible intensity	0.001 (0.015)	0.063 (0.041)	−0.007 (0.010)	−0.031** (0.014)
Export dummy	0.641*** (0.149)	0.085 (0.156)	0.318*** (0.097)	0.101** (0.045)
Year FE	Yes	Yes	Yes	Yes
Obs.	21,808	972	15,931	3,285
Panel B — Estimated ATEs:				
(1) Training within single-empl. firms vs. baseline: 0.206*** (0.020)				
(2) Multi-empl. firms vs. baseline: −0.021** (0.009)				
(3) Combined vs. baseline: 0.132*** (0.012)				
(4) Training vs. non-training, within multi-empl. firms: 0.153*** (0.013)				
(5) ATE(empl): −0.052** (0.024)				

Notes: The table reports second-step results and associated ATEs. The baseline group is “single-employee, no training” firms. Robust standard errors in parenthesis, clustered by firm, are obtained via firm-clustered bootstrap (200 replications). Asterisks denote significance levels (* at 10%, ** at 5% and *** at 1%).

– then the effect of training combined with scaling-up would likely differ systematically from the effect of training in single-employee firms.

More specifically, the analysis of the interplay of training with scaling-up from single-employee to multi-employee structure, can tell about two channels of particular interest, which we cannot directly measure in the data. First, to the extent that more complex micro-firms are also more complex in their HRM practices, then heterogeneity by firm size could also hint at the interplay of training with this otherwise completely omitted factor, crucial to the “bundle view” approach to HRM. Second, in single-employee firms we know with certainty who is receiving training, that is the owner/manager, a case often highlighted in the literature on SMEs as distinctively important.

With reference to the multi-valued IPWRA procedure described above, we take “being a multi-employee firm” as the focal dimension of interest and, thus, we define the following four groups: (i) the “baseline group” is made of single-employee firms which do not train; (ii) “training only” firms are single-employee firms which train; (iii) “dimension only” firms include multi-employee firms which do not train; (iv) the “combined” group encompasses multi-employee firms which spend in training.

Table 7 shows the IPWRA second-step results and the estimated ATEs, allowing to assess variation of the productivity return of training and the scaling-up/training complementarity.¹² Regarding the effect of training in single-employee firms – relative to the baseline of single-employee, no-training firms – the estimated ATE(training only) is positive (0.206) and highly significant, meaning that training raises productivity substantially without scaling-up. The ATE of scaling-up alone is negative (−0.021), telling that becoming multi-employee without training is actually associated with slightly lower productivity than staying single-employee and not investing in training. This is a small but statistically significant negative effect. The ATE of scaling-up together with investing in training (vis-a-vis the baseline) is positive (0.132) and highly significant, but notably smaller than the return of training in single-employee firms (0.206). Overall, the estimated combined effect (0.132) implies a shortfall of 0.052 (significant, p -value < 0.05) compared to an additivity benchmark for being jointly a “training

¹² Associated first-step results and covariate balance analysis are provided in Appendix B, see Table 14.

and multi-employee firm” of 0.185 (0.206–0.021). This significantly lower “combined ATE” is consistent with a story of substitutability, suggesting that organizational complexity does hamper the training-productivity pass-through. The productivity payoff of training is most cleanly realized when the single owner/manager directly absorbs and applies the new knowledge from training, while it is partly eroded when firms simultaneously expand their workforce, plausibly reflecting the costs of diffusing training benefits across, or coordinating them within, a larger team. Notwithstanding, this does not mean that training does not boosts productivity of multi-employee firms. Indeed, the estimated effect of moving from no-training to training within the group of multi-employee firms (0.153, strongly significant), reveals a still positive and significant productivity gain in the multi-employee group, but the magnitude is meaningfully smaller compared to the effect in the baseline group of single-employee firms (0.153 vs. 0.206). That is, overall, scaling-up does not annihilate the value of training, but rather the productivity-premium of training is smaller (of about 25%) once a firm passes the critical one-employee threshold.

5.2. Training and intangible investment

We then examine if the effect of training interacts with firms’ investment in intangible assets. This allows us to address interesting complementarities, which a-priori could or could not be in place according to different hypotheses. Under the resource-based/capability-based views of the firm and approaches to innovation which conceive training not merely as an accumulation of task-specific skills, training is itself a form of knowledge absorption and accumulation, and thus a crucial prerequisite, complementary factor to formal technological innovation activities (such as R&D, product/process innovation). In tune with this line of reasoning, one could expect that the productivity-effect of training might be amplified among those micro-firms which are already more dynamic and proactive in managing intangible capital. In fact, because the intangible investment variable recorded in our data encompasses R&D and IPR-related expenditures (in patents and trademarks) but it does not include training expenses – which other datasets may instead classify within intangibles too –, this variable is the closest proxy available to us for distinguishing micro-firms which are more likely to engage in formal innovation activities.¹³ Alternatively, following the mainstream approach to intangible capital of Corrado et al. (2005), in which training is itself conceived as a form of intangible investment, training and intangible investment recorded in our data may be expected to operate as substitute channels for accumulating knowledge capital, competing for the same productivity margin.

In the empirical analysis, we take as the focal dimension of heterogeneity whether firms have zero or positive intangible investments. This split is guided by the properties of the data. In fact, the median value of intangible investment in the working sample is zero, while only a small minority of firms record positive values (about 2% of the observations, recall Table 1). Accordingly, the multivalued IPWRA estimation is based on the following 4 groups: (i) zero-intangibles, no-training firms (the “baseline” reference group); (ii) zero-intangibles firms which train (“training only”); (iii) positive intangibles firms, which do not train (“dimension only”); and (iv) firms which invest in both training and intangibles (the “combined” group).

Table 8 shows the IPWRA second-step estimates and the associated ATEs.¹⁴ They show no evidence of statistically significant interaction

¹³ As per standard accounting principles complying with the Italian Civil Law (articles 2424 and 2435) – also valid in the case of the TREND dataset – recorded intangible assets essentially encompass R&D, IPRs (patents and trademarks, and related licenses/royalties) and goodwill not internally generated by the firm. But not training expenses.

¹⁴ Associated first-step results and covariate-balance analysis are provided in Appendix B, see Table 15.

Table 8

Training and intangibles — Multi-valued FE-IPWRA results.

Panel A — Second-step results:				
	No training, Intang. = 0	No training, Intang. > 0	Training, Intang. = 0	Training, Intang. > 0
Sales growth	0.019 (0.025)	0.017 (0.076)	0.026 (0.022)	0.027 (0.032)
Size	−0.363*** (0.051)	0.194 (0.392)	−0.765*** (0.063)	−0.834*** (0.486)
Tangible intensity	0.030*** (0.002)	−0.017 (0.039)	0.017*** (0.003)	0.033** (0.016)
Export dummy	0.389*** (0.070)	0.271 (0.560)	0.104** (0.044)	0.236 (0.167)
Year FE	Yes	Yes	Yes	Yes
Obs.	36,984	755	4,063	194
Panel B — Estimated ATEs:				
(1) No training and Intangibles > 0 vs. baseline: 0.138*** (0.052)				
(2) Training and Intangibles = 0 vs. baseline: 0.170*** (0.012)				
(3) Combined vs. baseline: 0.303*** (0.075)				
(4) Training vs. non-training, within positive intang. firms: 0.166* (0.095)				
(5) ATE(compl): −0.005 (0.096)				

Notes: The table reports second-step results and associated ATEs. Baseline group is firms which do not train and do not invest in intangibles. Robust standard errors in parenthesis, clustered by firm, obtained via firm-clustered bootstrap (200 replications). Asterisks denote significance levels (* at 10%, ** at 5% and *** at 1%).

between training and intangible investments. Compared to the baseline group, standalone investment in intangibles raises productivity by 0.138 and training alone by 0.170, while the estimated combined effect of jointly investing in both (0.303) essentially matches the simple sum of the separate effects (0.170+0.138 = 0.308). In fact, the resulting $ATE(compl)$ (−0.005) is far from significant, providing statistical support to an additivity result: the joint effect of training and intangibles equals the sum of their individual effects. This suggests that the two investments appear to operate as independent, separable productivity channels, each contributing its own effect, while combining them neither amplifies nor erodes the other's contribution. However, it is not just that the estimated ATE_{compl} is statistically zero. A further notable finding is that the effect of training within the group of firms that invest in intangibles, is numerically similar to the effect of training within firms that do not invest in intangibles: the two point estimates (0.170 vs. 0.166) are indeed almost identical. This is revealing that the training effect is essentially unchanged across the two groups of firms and, in turn, quite comparable to the aggregate ATE of training (0.177) obtained from the main analysis of Section 4.

5.3. Training and tangible investment

Next, we explore the interaction of training with investing in tangibles. The premise of this further heterogeneity analysis is that the joint effect of training and tangible investment on productivity is *a-priori* theoretically ambiguous. On the one hand, complementarity could arise in accordance to a story where firms may need training to effectively operate, maintain, and extract full productive value from machinery and equipment. This would imply that training is particularly valuable precisely when firms are also upgrading their physical capital, boosting the ability to reap productivity benefits from tangible investments. On the other hand, a certain degree of substitutability may also be hypothesized, as training and tangible investment could compete for the same underlying productivity margin, e.g., because both require managerial attention, organizational adjustment, or absorb the same pool of resources.

To assess between complementarity/substitutability in the multi-valued IPWRA, we take having zero or positive tangibles as the focal heterogeneity dimension. Similarly to the above analysis of intangible investment, the choice to dichotomize the sample across zero vs. positive tangible investment is guided by the properties of the

Table 9

Training and tangibles — Multi-valued FE-IPWRA results.

Panel A — Second-step results:				
	No training, Tangib. = 0	No training, Tangib. > 0	Training, Tangib. = 0	Training, Tangib. > 0
Sales growth	0.021** (0.009)	0.117*** (0.026)	0.020 (0.037)	0.036 (0.025)
Size	−0.330** (0.165)	−0.455*** (0.060)	−0.809*** (0.092)	−0.809*** (0.110)
Intangible intensity	0.026** (0.012)	−0.001 (0.009)	−0.028* (0.015)	0.015 (0.013)
Export dummy	0.471*** (0.110)	0.261*** (0.069)	0.097 (0.069)	0.059 (0.059)
Year FE	Yes	Yes	Yes	Yes
Obs.	27,091	10,644	2,188	2,068
Panel B — Estimated ATEs:				
(1) No training and Tangibles > 0 vs. baseline: 0.194*** (0.025)				
(2) Training and Tangibles = 0 vs. baseline: 0.192*** (0.025)				
(3) Combined vs. baseline: 0.295*** (0.029)				
(4) Training vs. non-training, within positive tangibles firms: 0.101*** (0.020)				
(5) ATE(compl): −0.091*** (0.033)				

Notes: Table reports second-step results and associated ATEs. Baseline group is firms which do not invest in either training or tangibles. Standard errors in parenthesis, clustered by firm, obtained via firm-clustered bootstrap (200 replications). Asterisks denote significance levels (* at 10%, ** at 5% and *** at 1%).

data: although investment in tangibles is not infrequent (about 30% of the sample record non-zero tangibles, recall Table 1), the median investment in tangibles in the working sample is actually zero. Accordingly, the multi-valued IPWRA analysis is based on the following four mutually exclusive groups: (i) firms with neither training nor tangible investment (“baseline” group), (ii) firms which invest in tangibles, but do not train (“dimension only”); (iii) firms which train, but do not invest tangibles (“training only”); and (iv) firms which invest in both training and tangibles (the “combined” group).

The second-step results and associated ATEs, presented in Table 9, support substitutability.¹⁵ Both training and tangible investments, when undertaken on their own, raise productivity substantially compared to the baseline: tangible investments alone yield an ATE of 0.194, while training alone yields a nearly identical ATE of 0.192 (both strongly significant). However, the estimated combined effect is 0.295, which implies a shortfall of 0.091 (see $ATE(compl)$, highly statistically significant) compared to the additivity benchmark of 0.386 (0.192+0.194) that would arise by simply summing the two separate effects.

This substitutability result, implying that the incremental contribution of training is comparatively reduced if physical capital investment is also in place, is consistent with the notion that firms may face coordination or absorption costs when upgrading physical capital and human capital simultaneously. Also, it suggests that training may be a particularly powerful lever for boosting productivity precisely among micro-firms which do not simultaneously devote resources to tangibles, and whose productivity is therefore very much dependent on knowledge upgrading of the employees. In fact, while training still remains a positive contributor to productivity irrespective of whether it is coupled to tangibles, the productivity gain of training estimated within the firms which do not invest in tangibles (0.192) is roughly double compared to that estimated for firms which invest in tangibles (0.101, highly significant).

5.4. Training and sectoral specificity of manufacturing

Lastly, we examine the possible role of sectoral specificities, considering that the aggregate baseline estimates may hinder relevant heterogeneities in the training–productivity relationship if training needs,

¹⁵ Associated first-step results and covariate-balance analysis are provided in Appendix B, see Table 16.

Table 10
Training and manufacturing — Multi-valued FE-PWRA results.

Panel A — Second-step results:				
	Non-Manuf., No training	Non-manuf., training	Manuf., No training	Manuf., training
Size	−0.345*** (0.058)	−0.745*** (0.065)	−0.278** (0.116)	−1.079*** (0.300)
Sales growth	0.027 (0.032)	0.021 (0.015)	0.018 (0.029)	0.079 (0.113)
Tangible intensity	0.028*** (0.002)	0.017*** (0.003)	0.034*** (0.005)	0.028* (0.016)
Intangible intensity	−0.009 (0.009)	−0.004 (0.010)	0.012 (0.025)	−0.074* (0.044)
Export dummy	0.348*** (0.069)	0.089* (0.051)	0.492*** (0.171)	0.219 (0.152)
Year FE	Yes	Yes	Yes	Yes
Obs.	30,670	3,585	7,069	672
Panel B — Estimated ATEs:				
(1) Training in non-manufacturing firms vs. baseline: 0.172*** (0.013)				
(2) Manufacturing, no training firms vs. baseline: −0.007 (0.009)				
(3) Combined vs. baseline: 0.164*** (0.031)				
(4) Training vs. non-training within manufacturing: 0.170*** (0.032)				
(5) ATE(compl): −0.002 (0.034)				

Notes: The table reports second-step results and associated ATEs. Baseline group is firms which are outside manufacturing and do not spend in training. Standard errors in parenthesis, clustered by firm, obtained via firm-clustered bootstrap (200 replications). Asterisks denote significance levels (* at 10%, ** at 5%, and *** at 1%).

content or applicability differ systematically across sectors. The TREND dataset reports information of 4 “macro-sectors” (manufacturing, construction, trade, and services), thus preventing to provide a detailed analysis by narrowly defined industries and to apply taxonomies related to knowledge dynamics (e.g., by Pavitt taxonomy classes or between high-tech vs. low-tech tech industries). We focus on a contrast between manufacturing vs. non-manufacturing firms, since manufacturing is ex-ante the macro-sector most likely to exhibit complementarities with training, given its greater reliance on technical, machine- and process-specific knowledge and competences that training might directly affect.

Accordingly, in the multivalued FE-IPWRA analysis, we define the following four groups, taking membership in manufacturing as the focal dimension “to pair” with training: (i) non-manufacturing, no-training firms (“baseline”); (ii) non-manufacturing firms which train (“training only”); (iii) manufacturing firms which do not train (“dimension only”); and (iv) manufacturing firms which invest in training (“combined” category).

Table 10 reports the second-step regression results and the estimated ATEs.¹⁶ When considering the separate effects, they show that, among non-training firms, being active in manufacturing does not associate with a productivity-premium vis-a-vis non-manufacturing firms (ATE is −0.007, not significant), whereas investing in training among non-manufacturing firms raises productivity by 0.172. If the additive vs. combined effects are considered, we find no significant interaction with training. The joint effect (0.164) closely matches the additive benchmark (0.172 − 0.007 = 0.165), and indeed the *ATE(compl)* (−0.002) is statistically equal to zero. This additivity result indicates that the productivity payoff of training is sector-neutral, or at least it does not appear to be disproportionately activated by the more technically-intensive character of manufacturing firms, ruling out this channel as a major driver of the aggregate result. In fact, we also notice that training effects have the same (positive and significant) magnitude, regardless of whether the firm operates in manufacturing: the estimated effect of moving from no-training to training within

manufacturing (0.170) is essentially identical to the ATE of training within the baseline group of non-manufacturing firms (0.172). And they are substantially identical to the aggregate ATE of training (0.177) estimated in Section 4.

To sum up, the heterogeneity analysis, read together, delivers two main patterns. First, the estimated ATEs across all specifications, all reveal a positive, statistically significant and economically sizable effect of training. On the one hand, the “standalone” effects of training within the various baseline groups (single-employee, zero tangibles or intangibles, non-manufacturing), are remarkably stable across specifications: the estimated ATEs range from 0.170 to 0.206, fall within one standard error of each other in every pairing and all closely track (within one standard error) the aggregate FE-IPWRA estimate of the ATE of training (0.177). On the other hand, the ATEs of training within the non-baseline groups (multi-employee, positive tangibles or intangibles, manufacturing), are also comparable in magnitude – ranging in between 0.101 and 0.194 – to the aggregate estimate of the ATE of training, although they are statistically smaller in two cases (0.153 for multi-employee firms and 0.101 for positive tangibles firms). This broad consistency indicates that a positive productivity-premium of training is not an artifact of any particular specification or comparison group, but reflects a robust underlying relationship between training and productivity. This strongly corroborates the conclusion from the aggregate analysis that training provides a positive contribution to productivity of micro-firms.

Second, despite the positive effect of training is not tightly dependent from any of the considered heterogeneity dimensions, we document that the marginal value of training is not affected by co-occurrence with intangible investment or by the firm’s manufacturing-sector membership, while measurably reduced when training co-occurs with tangibles-expanding or scale-expanding investments. In both cases, the data reject additivity in favor of substitutability, implying that the decision to invest in training appears as more effective, comparatively, in boosting productivity in single-employee firms and in firms which, by not investing in tangibles, rely upon their workforce knowledge accumulation as a source of competitiveness. The estimated contribution of training is particularly strong in this latter case.

6. Conclusions

This paper contributes to the literature on the training–productivity nexus by focusing on micro-firms, a heavily under-researched category. The common wisdom that training enhances productivity – based on studies of SMEs and large firms – may not hold for micro-firms. They exhibit distinctive characteristics, even against small-medium firms, and cannot be treated as a scale-down of larger enterprises.

On the one hand, because people (owners/managers and a few employees) are central in micro-firms, training could represent a crucial channel for enhancing productivity, as a direct mode of accumulation and absorption of knowledge available to this category of firms, which in fact typically lack organizational or financial resources to engage in more structured innovation strategies that may spur competitiveness –involving, e.g. R&D, patents or process/product innovation. However, on the other hand, whether the training–productivity channel positively operate in practice is not obvious. Knowledge accumulation via training – whether it works as a complementary precondition for or as an alternative route to technological innovation – does not need to automatically pass-through productivity improvements, as it may prove effective only conditional upon certain firm-specific characteristics. Also, in the specific context of micro-firms, training can divert scarce time and financial resources without delivering sufficient productivity gains, while competence and knowledge often develop informally, on-the-job and through learning mechanisms other than training. And, finally, micro-firms may end up with no productivity return for training as they may see their trained employees poached by larger firms.

¹⁶ Associated first-step coefficient estimates and covariate-balance analysis are provided in Appendix B, see Table 17.

These considerations, coming from various theoretical underpinnings, suggest that whether training can raise productivity in micro-firms is ultimately an empirical question. Yet, existing studies do not challenge the conventional wisdom that training fosters productivity in micro-firms as it does in firms of larger size. We address this research gap by estimating whether micro-firms that invest in training enjoy a productivity premium compared to micro-firms that do not. The empirical analysis relies on a unique dataset reporting training expenditures for a large, representative panel of Italian micro-firms.

Our main result is to document a clear positive productivity premium: on average, training firms are significantly more productive than non-training firms. The estimated ATE of training, controlling for endogeneity of training status via IPWRA methods is sizable – about 19% in the more reliable estimates including firm fixed-effects. The estimated effect remains positive and overall comparable in magnitude across the dimensions of heterogeneity which our data allows us to explore, namely across micro-firms of different size (single-employee vs. other micro-firms), across micro-firms active in manufacturing vs. other sectors, and no matter whether they invest in intangible or tangible assets. This stability of results supports that our main finding reflects a robust underlying relationship between training and productivity, not tightly dependent from any of the considered heterogeneity dimensions.

Since both the treated group of training firms and the control group of non-training firms are likely to rely on informal learning mechanisms other than training (e.g., on-the-job coaching) – we interpret the observed productivity premium as suggesting that training enhances productivity above and beyond the alternative modes of human capital accumulation typical of micro-firms. In this respect, our findings provide support to policies that encourage training investment by micro-firms, as there is hope to boost productivity of micro-firms beyond the productivity levels that can be reached through learning modes already in place in micro-firms. A number of existing SMEs-oriented policies could be re-targeted or better tailored to micro-firms, from awareness campaigns addressing ignorance-related biases, to actions aiming to reduce market failures such as financial constraints or fears of poaching. While we do not directly test in our data whether training effects are stronger in micro-firms than in larger firms, the sizable estimated effect supports that well-designed interventions could help narrowing the well-known productivity gap between micro and larger firms.

An interesting qualification coming from the heterogeneity analysis, is that the contribution of training – while essentially neutral to manufacturing sector membership and co-occurrence of intangible investments – could be sensibly reduced for firms that combine training with investing in tangibles or with scaling-up from a single-employee to a multi-employee structure. On the one hand, the neutrality result with respect to the joint presence of intangible investments in the firm, is particularly interesting for innovation studies, since it contradicts the hypothesis that training would be effective only if coupled with engagement in technological innovation or formal innovation activities, such as R&D or patenting, of which intangibles provide at least an indirect proxy in our data. The two investments appear to operate as independent, separable productivity channels, each contributing its own effect, while combining them neither amplifies nor erodes the other's contribution. On the other hand, the reduced marginal effect of training estimated among firms that combine training with tangible capital deepening or workforce expansion – compatible with a substitutability story – suggests that coordination costs, absorption constraints, or organizational adjustment associated with physical capital or labor expansion could partially erode the incremental value of training. The reduced marginal contribution of training is particularly remarkable in micro-firms that invest in tangibles, in line with the intuition that training matters most for micro-firms which, by not investing in tangible capital, rely on knowledge accumulation of employees as their primary source of productivity.

We recognize several limitations of our study, which may in turn open up valuable avenues for future research on the training-productivity link in micro-firms. These revolve around the general and well-known challenge of accessing detailed data on micro-firms. In fact, despite we could exploit the unique opportunity to measure training expenditures for a large set of micro-firms, improving the availability, coverage and quality of data still remains a priority for future development of micro-business research, even beyond the specific topic of our study.

At a general level, we urge for extending initiatives like the TREND project. Wider access to datasets comparable to TREND, allowing to link information on training with data about firms' structure and performance, for large panels of micro-firms, would prove essential to assess whether our findings are specific to our context or can be generalized, primarily across other Italian regions but also across other countries. In this respect, we offer a starting point for future research that could confirm external validity beyond the context of our dataset.

More specifically, we could not capture many interesting factors, potentially relevant for a deeper understanding of the workings of the training-productivity link, due to lacking appropriate data. The variables available to us, allowed to achieve covariate balance and reasonable confidence in the identification of the training effect, and we could also explore substitutability/complementarity effects between training and interesting dimensions of heterogeneity. But future studies could surely dig more thoroughly into the mechanisms of the training-productivity link, by accessing more and more detailed information on various dimensions.

A first promising direction that we foresee, concerns collecting data on a number of attributes internal to the firm, which could help to better characterize the “unfolding of training within firms”. In this category of additional data, the dimension of perhaps more direct interest for innovation studies, very much related to our heterogeneity analysis, regards the overall innovation strategy of firms. While the interpretation of intangible investment as an indirect, loose measure of engagement in innovation activity of firms can be to some extent accepted, one would like to have more detailed data, primarily about standard measures of technological innovation such as R&D, patenting or product/process innovation. Because our data do not include these measures, we cannot directly test whether the estimated productivity premium of training operates through an innovation channel (e.g., training enabling absorption of external knowledge or supporting incremental process innovation) or through a pure human-capital/efficiency channel unrelated to innovation. Two conjectures, motivated by our heterogeneity results, could guide future work with richer data: (i) if the effect of training on productivity is mediated by innovation, we would expect the training premium to be larger, not statistically indistinguishable, among firms with positive R&D or patenting activity – contrary to what our loose proxy (investment in intangibles) has suggested; (ii) if training substitutes for innovation as a competitiveness strategy for capital-light micro-firms, as suggested by our finding of a larger premium among firms with zero tangible investments, we would expect training-only firms to show innovation outcomes (e.g., informal or non-patented process improvements) that are under-counted in standard innovation surveys.

Always within the category of data which could shed light on the “internal unfolding of training”, another dimension which would be worth trying and consider in future studies, concerns the personal profiles (age, educational background, previous experience or social networks) of the core people in the firm (owner/manager and other employees), in particular distinguishing who is exposed to training from who is not. These features – to our knowledge – are only measured in limited in scope, ad-hoc surveys covering few firms, or in specific firm case-studies. Their coverage should be extended to larger datasets. Our analysis across single-employee vs. multi-employee firms supports that training effects seem stronger when the owner/manager is the person who receives and applies training, while returns are reduced in more complex organizations. Yet, with more detailed data

one could: (i) explore more deeply the role played by the personal profiles of the core people in the firm, particularly of those receiving training; (ii) address the question whether the estimated training effect works via individuals' upgrading (in line with the neoclassical theory) vs. organizational-level knowledge accumulation and absorptive capacity, a distinction central to the innovation literature but currently untestable in our dataset; (iii) test the poaching risk hypothesis – to the extent that one can also access data on job-flows via matched employer–employee data – by directly examining if trained people leaves the firm and when.

Relatedly, a further crucial dimension for a deeper understanding of how training works within firms, of particular interest for managerial studies and for which having more detailed data would be surely important, is represented by information about HRM practices in place in the firms. While the lack of this type of information is not just specific to our dataset (again, only small ad-hoc surveys usually have that), our study can only indirectly tell something about its role, based on the comparison between single-employee vs. multi-employee firms. Assuming that scaling-up in size goes together with adopting more complex HRM practices, our conjecture, based on the finding that multi-employee firms see their productivity return to training reduced, is that training may not necessarily have beneficial effects if not well integrated in the design of complex HRM practices. But more detailed data, possibly for large samples of micro-firms, are needed to explicitly address the non-trivial integration of training with other elements of the HRM strategy, as theorized in the “bundle view” approach to HRM in micro-firms.

Second, complementary to expanding data about the unfolding of training within firms, future studies would benefit from digging deeper into the heterogeneity across firms. Particularly through expanding the data to better measure different micro-firms typologies. An easy example is that, by integrating reliable data about foundation year, one could examine variation of training effects across older established micro-firms vis-a-vis very young, micro start-ups. More generally, one would like to account for the fact that the very notion of micro-businesses encompasses a range of heterogeneous entities, embracing not only start-ups, but also other quite diverse types such as owner-managed vs. team-managed businesses, family businesses, ethnic businesses, high-tech businesses and businesses with differing legal status. As these different typologies of micro-firms are known to be characterized by peculiar learning dynamics (Devins et al., 2005), the productivity-training linkages may also differ across them. One could replicate our approach across such diverse entities. That could have important implications, in particular, for informing the design of policies, which, at present, tend to treat the category of micro-firms as a “one-size fits all” notion. Or, even more unfortunately, as entirely overlapping with the broader category of SMEs.

Third, beyond more detailed data about within-firm characteristics and firm types, there is also a need to improve upon the measurement of the characteristics and the content of training programs that firms invest into. Features such as in-house vs. ex-house training or the degree of standardization vs. tailoring with firm-specific needs, are part of the discussion in the SMEs literature (Bryan, 2006; Patton et al., 2000). They may all play a role in whether and how training investment successfully contributes to productivity of micro-firms. In fact, some types of training may arguably be more relevant to productivity – such as learning about digital systems, new software or machinery – while other contents of training – such as leadership development, fair and equitable working environment, or compliance with regulations – may have more uncertain productivity returns. Once again, data about these elements are usually available in small/targeted surveys or explored in depth only in specific firm case studies. It would be interesting to test their use and role in more comprehensive datasets of micro-firms.

Despite its limitations, which point to interesting avenues for future research, our study is, to our knowledge, the first attempt to assess the effect of training on productivity in micro-firms, exploiting a large

sample and more solid empirical methods than are usually employed in the micro-firms literature. We hope it could serve as a benchmark for future research and encourage greater attention to this important issue. More, and more robust, evidence is needed on the drivers of productivity in micro-businesses, given their predominance in every economy.

CRediT authorship contribution statement

Marco Rispoli: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Federico Tamagni:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has been developed within a research collaboration between ISTAT, CNA and Scuola Superiore Sant'Anna, while one of the authors (Marco Rispoli) was visiting ISTAT. We want to thank all the people involved in accessing and managing the data, in particular Tommaso Rondinella, Luca Faustini, Roberta Palmieri and the researcher at ISTAT facilities in Florence. Furthermore, we would like to thank the Presidents and the staff of the regional CNA offices of Marche, Umbria and Toscana, who authorized to use the microdata, under the mandatory condition of censorship of individual information.

Appendix A. Robustness to alternative estimators

As mentioned in the main text, IPWRA estimates are theoretically preferred to IPW or RA methods since IPWRA is a doubly-robust methodology: while RA performs poorly when the outcome model is mis-specified and IPW performs poorly under mis-specification of the treatment model, IPWRA only requires that either the outcome or the treatment model is well specified. However, since model mis-specification is untestable, there remains the possibility that both models are mis-specified in a given empirical application. In such a situation, which estimator is more reliable is a matter of debate in the literature. The suspect is that the answer depends on the true models, the implemented specifications and the set of covariates used. But there is no definitive conclusion in the literature.

In our analysis, despite we achieve covariate balance and apply a variant of the IPWRA including firm fixed-effects, mis-specification of both outcome and treatment model is a possibility. It is then worth checking whether the main effects are sensitive to changing estimation method. We here report IPW and RA estimates, and also the results obtained via an additional estimation technique, the Augmented IPW (AIPW) estimator. This method is doubly-robust and thus theoretically similar to the IPWRA estimator. However, no literature compares the relative performance of the AIPW estimator, which emerged from a general approach to creating efficient estimators, against the IPWRA estimator, which emerged from a robust correction approach to missing-data analysis.

In Table 11 we display OLS estimates of the outcome models and the corresponding estimated ATE of training across the different methodologies. We also include the OLS-IPWRA estimates already reported

Table 11
Alternative estimation methods — Outcome models, OLS results and associated ATEs.

	OLS-IPW	OLS-RA		OLS-IPWRA		OLS-AIPW	
	(1) ln Prod	(2) ln Prod ⁰	(3) ln Prod ¹	(4) ln Prod ⁰	(5) ln Prod ¹	(6) ln Prod ⁰	(7) ln Prod ¹
Sales growth	–	0.022*** (0.009)	0.004** (0.002)	0.020*** (0.007)	0.011** (0.005)	0.022*** (0.009)	0.004** (0.002)
ln(Size)	–	0.049*** (0.016)	–0.025 (0.020)	0.050*** (0.016)	–0.017 (0.024)	0.049*** (0.016)	–0.025 (0.020)
ln(Tangible intensity)	–	0.109*** (0.002)	0.057*** (0.004)	0.108*** (0.003)	0.058*** (0.004)	0.109*** (0.002)	0.057*** (0.004)
ln(Intangible intensity)	–	–0.0004 (0.012)	0.018** (0.009)	–0.003 (0.011)	0.037*** (0.013)	–0.004 (0.012)	0.018** (0.009)
Export dummy	–	0.767*** (0.038)	0.360*** (0.063)	0.755*** (0.038)	0.402*** (0.083)	0.767*** (0.038)	0.360*** (0.063)
Obs.	42,658	38,329	4,329	38,329	4,329	38,329	4,329
ATE	0.594*** (0.023)	0.622*** (0.021)		0.614*** (0.024)		0.614*** (0.023)	

Notes: The table reports OLS estimates of the outcome models (second-step weighted regressions), and the corresponding ATE of training across different methodologies (IPW, RA, IPWRA and AIPW). Regressions also include year, region and sector fixed-effects. Robust standard errors in parenthesis, clustered by firm and corrected via the STATA *teffects* command to account for the two step procedure. Asterisks denote significance levels (* at 10%, ** at 5% and *** at 1%).

Table 12
Alternative estimation methods — Outcome model, FE results and associated ATEs.

	FE-IPW	FE-RA		FE-IPWRA		FE-AIPW	
	(1) ln Prod	(2) ln Prod ⁰	(3) ln Prod ¹	(4) ln Prod ⁰	(5) ln Prod ¹	(6) ln Prod ⁰	(7) ln Prod ¹
Sales growth	–	0.025*** (0.008)	0.015 (0.014)	0.023*** (0.007)	0.021 (0.015)	0.025*** (0.008)	0.015 (0.014)
ln(Size)	–	–0.301*** (0.046)	–0.698*** (0.048)	–0.302*** (0.045)	–0.805*** (0.064)	–0.301*** (0.046)	–0.698*** (0.048)
ln(Tangible intensity)	–	0.029*** (0.002)	0.015*** (0.003)	0.030*** (0.002)	0.0180*** (0.004)	0.029*** (0.003)	0.015*** (0.002)
ln(Intangible intensity)	–	–0.002 (0.008)	–0.011 (0.009)	–0.003 (0.008)	–0.012 (0.012)	–0.002 (0.008)	–0.011 (0.009)
Export dummy	–	0.438*** (0.065)	0.082** (0.037)	0.418*** (0.062)	0.100** (0.045)	0.438*** (0.065)	0.082*** (0.037)
Obs.	42,658	38,329	4,329	38,329	4,329	38,329	4,329
ATE	0.168*** (0.013)	0.178*** (0.009)		0.177*** (0.011)		0.174*** (0.011)	

Notes: The table reports firm fixed-effects (FE) estimates of the outcome models (second-step weighted regressions), and the corresponding ATE of training across different methodologies (IPW, RA, IPWRA and AIPW). Corrected standard errors, clustered at the firm level, are obtained via bootstrap (200 replications). Asterisks denote significance levels (* at 10%, ** at 5% and *** at 1%).

in the main text, in order to ease comparisons.¹⁷ The results are very similar across methods. In particular, the estimated ATEs of training are stable across all specifications in terms of sign, magnitude and significance. In fact, the point estimates (0.594, 0.622, 0.614 and 0.614) are all statistically equal within a 1-standard error confidence band. This lend support to the conclusions from the OLS-IPWRA analysis presented in the main text.

In Table 12 we present the estimates with firm fixed-effects, again reporting our baseline FE-IPWRA for ease of comparison to the main text.¹⁸ Also here we observe that results do not change significantly across the estimation methods. The ATEs estimated via RA and AIPW are quite in line with the baseline FE-IPWRA results, and actually equal to the latter within a 1-standard error confidence band. Overall, they all support a positive and significant productivity premium of training,

whose magnitude ranges in between 18.3% (0.168 log-points) and 19.5% (0.178 log-points).

Appendix B. Robustness and heterogeneity analysis: IPWRA first-step results and covariate balance

To ease readability of the main text, we document here about all the IPWRA first-step estimates and the covariate-balance diagnostics associated to robustness checks and the heterogeneity analysis.

Table 13 refers to the robustness check with respect to lagged training status, performed in Section 4. It reports number of observations for the unweighted and re-weighted samples, first-step Probit estimates, and covariate-balance diagnostics.

Regarding the heterogeneity analysis of Section 5, based on the modified multi-valued FE-IPWRA framework, we report number of observations for the unweighted and re-weighted samples, the first-step Multinomial Logit results (with the reference omitted group given by the various baseline categories), and the associated covariate-balance diagnostics. Table 14 refers to the heterogeneity analysis by single- vs. multi-employees firms. Table 15 refers to the estimates across zero vs. positive intangible investments. Table 16 reports about the analysis by zero vs. positive investment in tangibles. Table 17 shows the results of the analysis by manufacturing vs. other sectors.

¹⁷ Coefficient estimates and standard errors are obtained via the STATA command *teffects*, which implement appropriately corrected standard errors. Notice that first-step Probit estimates of the p-scores are invariant across the methods. The associated coefficient estimates and covariate-balance analysis are exactly identical to those discussed in the main text.

¹⁸ Since the correction for standard errors to account for the two-step procedure is not available via the STATA *teffects* command in the case of firm FE models, the corrected standard errors in the firm FE estimates were obtained via bootstrap.

Table 13
Lagged (1-year) training status — IPWRA first-step results and covariate balance.

Observations	Raw	Weighted			
Total Obs.	25,568	25,568			
– Treated Obs.	2,714	12,767.9			
– Control Obs.	22,854	12,800.1			
Variable	Probit Coeffs	Raw Std diff	Weighted Std diff	Raw var ratio	Weighted var ratio
Sales growth _{<i>t-1</i>}	–0.002 (0.002)	–0.003	–0.010	0.543	0.523
Single-empl. dummy _{<i>t-1</i>}	–0.678*** (0.026)	–0.718	–0.004	0.721	1.000
Tangible dummy _{<i>t-1</i>}	0.261*** (0.024)	0.379	0.008	1.152	1.007
Intangible dummy _{<i>t-1</i>}	0.131** (0.062)	0.139	0.003	2.106	1.022
Export dummy _{<i>t-1</i>}	0.062 (0.067)	0.070	–0.008	1.492	0.952
Covariate-balance test:	Test stat = 3.010 p-value = 0.8076				

Notes: The table reports number of observations in the raw and re-weighted sample, together with first-step Probit coefficients and covariate-balance diagnostics (standardized differences and variance ratios). Cluster-robust (by firm) standard errors in parentheses, asterisks denote significance levels (* at 10%, ** at 5%, and *** at 1%). Covariate-balance test is the [Imai and Ratkovic \(2014\)](#) over-identifying restriction test for covariate balance: χ^2 test statistic and corresponding *p*-value are reported.

Table 14
Heterogeneity by firm size — First-stage results and covariate balance.

Panel A — First-stage Multinomial Logit:					
	Single-empl., train vs. baseline	Multi-empl., no train vs. baseline	Multi-empl., train vs. baseline		
Sales growth	−0.001 (0.005)	0.004 (0.003)	0.003 (0.003)		
Tangible dummy	0.734*** (0.074)	0.876*** (0.028)	1.409*** (0.045)		
Intangible dummy	0.317 (0.274)	1.188*** (0.097)	1.386*** (0.119)		
Export dummy	−0.236 (0.338)	0.651*** (0.109)	0.760*** (0.149)		
Panel B — Covariate-balance diagnostics:					
Obs:	Baseline	Single-empl., train	Multi-empl., no train	Multi-empl., train	
Raw	21,808	972	15,931	3,285	
Weighted	10,497.5	10,499.8	10,503.3	10,495.3	
		Standardized differences		Variance ratio	
		Raw	Weighted	Raw	Weighted
<i>Single-empl., train:</i>					
Sales growth		0.012	−0.014	0.190	0.637
Tangible dummy		0.328	0.002	1.406	1.001
Intangible dummy		0.064	0.003	1.859	1.021
Export dummy		−0.014	−0.008	0.891	0.949
<i>Multi-empl., no train:</i>					
Sales growth		0.026	0.005	7.138	1.212
Tangible dummy		0.424	0.001	1.477	1.001
Intangible dummy		0.198	0.009	4.643	1.059
Export dummy		0.116	−0.002	2.202	0.989
<i>Multi-empl. and train:</i>					
Sales growth		0.040	0.002	3.551	1.063
Tangible dummy		0.717	0.002	1.546	1.002
Intangible dummy		0.273	0.011	6.777	1.073
Export dummy		0.154	0.001	2.693	1.004

Notes: Panel A reports the IPWRA first-stage Multinomial Logit coefficients. Baseline omitted category is “single-employee, non-training firms”. Robust standard errors in parentheses, clustered by firm. Asterisks denote significance levels (* at 10%, ** at 5%, and *** at 1%). Panel B reports raw and IPW-weighted observations by group, together with standardized differences and variance ratios for each covariate, relative to the baseline category.

Table 15
Heterogeneity by intangibles — First-stage results and covariate balance.

Panel A — First-stage Multinomial Logit:					
	Intang. > 0, no train vs. baseline	Intang. = 0, train vs. baseline	Intang. > 0 and train vs. baseline		
Sales growth	0.005 (0.004)	−0.004 (0.004)	0.007 (0.005)		
Dummy individual	−1.195*** (0.097)	−1.406*** (0.046)	−2.420*** (0.284)		
Tangible dummy	1.898*** (0.091)	0.583*** (0.037)	2.398*** (0.209)		
Export dummy	0.784*** (0.187)	0.082 (0.117)	0.801*** (0.294)		
Panel B — Covariate-balance diagnostics:					
Obs:	Baseline	Intang. > 0, no train	Intang. = 0, train	Intang. > 0 and train	
Raw	36,984	755	4,063	194	
Weighted	10,808.4	11,347.5	10,744.6	9095.5	
		Standardized differences		Variance ratio	
		Raw	Weighted	Raw	Weighted
<i>Intang. > 0, no train:</i>					
Sales growth	0.060	0.015	9.402	0.834	
Dummy individual	−0.795	0.088	0.714	0.977	
Tangible dummy	1.131	−0.046	0.911	0.958	
Export dummy	0.238	−0.028	3.245	0.826	
<i>Intang. = 0, train:</i>					
Sales growth	0.005	−0.011	0.283	0.514	
Dummy individual	−0.760	−0.009	0.742	1.001	
Tangible dummy	0.414	0.020	1.257	1.017	
Export dummy	0.064	−0.007	1.484	0.957	
<i>Intang. > 0, train:</i>					
Sales growth	0.117	0.029	13.915	0.899	
Dummy individual	−1.303	−0.203	0.277	0.993	
Tangible dummy	1.475	0.170	0.608	1.120	
Export dummy	0.267	0.003	3.624	1.019	

Notes: Panel A reports the IPWRA first-stage Multinomial Logit coefficients. Baseline omitted category is “zero intangibles, non-training firms”. Robust standard errors in parentheses, clustered by firm. Asterisks denote significance levels (* at 10%, ** at 5%, and *** at 1%). Panel B reports raw and IPW-weighted observations by group, together with standardized differences and variance ratios for each covariate, relative to the baseline category.

Table 16
Heterogeneity by tangibles — First-stage results and covariate balance.

Panel A — First-stage Multinomial Logit:				
	Tang. > 0, no train vs. baseline	Tang. = 0, train vs. baseline	Tang. > 0 and train vs. baseline	
Sales growth	0.039** (0.017)	−0.025 (0.023)	0.042** (0.018)	
Dummy individual	−0.876*** (0.028)	−1.476*** (0.055)	−2.154*** (0.070)	
Intangible dummy	1.897*** (0.092)	0.267 (0.208)	2.091*** (0.118)	
Export dummy	0.666*** (0.086)	0.139 (0.173)	0.692*** (0.148)	
Panel B — Covariate-balance diagnostics:				
Obs:	Baseline	Tang. > 0, no train	Tang. = 0, train	Tang. > 0 and train
Raw	27,091	10,644	2,188	2,068
Weighted	10,585.2	10,472.7	10,488.1	10,445.0
		Standardized differences		Variance ratio
		Raw	Weighted	Raw Weighted
<i>Tang. > 0, no train:</i>				
Sales growth	0.076	−0.124	1.249	0.717
Dummy individual	−0.469	0.003	1.056	1.000
Intangible dummy	0.280	0.008	7.842	1.055
Export dummy	0.131	0.002	2.294	1.014
<i>Tang. = 0, train:</i>				
Sales growth	−0.011	−0.138	0.118	0.806
Dummy individual	−0.754	0.002	0.897	1.000
Intangible dummy	0.060	0.005	1.868	1.031
Export dummy	0.049	−0.007	1.420	0.953
<i>Tang. > 0, train:</i>				
Sales growth	0.084	−0.120	1.694	0.815
Dummy individual	−1.120	0.000	0.594	1.000
Intangible dummy	0.368	0.014	11.316	1.094
Export dummy	0.167	−0.008	2.754	0.948

Notes: Panel A reports the IPWRA first-stage Multinomial Logit coefficients. Baseline omitted category is “zero tangibles, non-training firms”. Robust standard errors in parentheses, clustered by firm. Asterisks denote significance levels (* at 10%, ** at 5%, and *** at 1%). Panel B reports raw and IPW-weighted observations by group, together with standardized differences and variance ratios for each covariate, relative to the baseline category.

Data availability

The authors do not have permission to share data.

References

- Acemoglu, D., Pischke, J.-S., 1998. Why do firms train? Theory and evidence. *The Q. J. Econ.* 113 (1), 79–119.
- Acemoglu, D., Pischke, J.-S., 1999. Beyond becker: Training in imperfect labour markets. *The Econ. J.* 109 (453), 112–142.
- Ao, W., Calónico, S., Lee, Y.-J., 2021. Multivalued treatments and decomposition analysis: An application to the WIA program. *J. Bus. Econom. Statist.* 39 (2), 358–371.
- Barney, J., 1991. Firm resources and sustained competitive advantage. *J. Manage.* 17 (1), 99–120.
- Bartel, A.P., 1994. Productivity gains from the implementation of employee training programs. *Ind. Relat.: A J. Econ. Soc.* 33 (4), 411–425.
- Baumann, J., Kritikos, A.S., 2016. The link between R&D, innovation and productivity: Are micro firms different? *Res. Policy* 45 (6), 1263–1274.
- Becker, G.S., 1964. Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition. NBER Books, National Bureau of Economic Research, Inc.
- Berlingieri, G., Blanchenay, P., Calligaris, S., Criscuolo, C., 2017. The Multiprod Project: A Comprehensive Overview. Technology and Industry Working Papers 2017/04, OECD Publishing.
- Black, S.E., Lynch, L.M., 1996. Human-capital investments and productivity. *The Am. Econ. Rev.* 86 (2), 263–267.
- Brancati, E., Brancati, R., Guarascio, D., Zanfei, A., 2022. Innovation drivers of external competitiveness in the great recession. *Small Bus. Econ.* 58 (3), 1497–1516.

Table 17
Heterogeneity by manufacturing firms — First-stage and covariate balance.

Panel A — First-stage Multinomial Logit:				
	No manufact., train vs. baseline	Manufact., no train vs. baseline	Manufact., train vs. baseline	
Sales growth	−0.001 (0.002)	−0.004 (0.006)	−0.008 (0.012)	
Dummy individual	−1.393*** (0.049)	−0.414*** (0.043)	−2.090*** (0.146)	
Tangible dummy	0.573*** (0.041)	−0.005 (0.037)	0.616*** (0.088)	
Intangible dummy	0.290*** (0.101)	0.578*** (0.095)	0.618*** (0.174)	
Export dummy	0.121 (0.156)	1.375*** (0.104)	1.642*** (0.183)	
Panel B — Covariate-balance diagnostics:				
Obs:	No manuf., no train	No manuf., train	Manuf., no train	Manuf., train
Raw	30,670	3,585	7,069	672
Weighted	10,529.7	10,468.1	10,533.7	10,464.5
		Standardized differences		Variance ratio
		Raw	Weighted	Raw Weighted
<i>No manuf., train:</i>				
Sales growth	0.017	−0.002	0.898	0.844
Dummy individual	−0.765	−0.009	0.773	1.002
Tangible dummy	0.426	0.018	1.249	1.015
Intangible dummy	0.148	0.004	2.440	1.028
Export dummy	0.057	0.000	1.540	0.998
<i>Manuf., no train:</i>				
Sales growth	−0.005	0.003	1.161	1.149
Dummy individual	−0.229	0.000	1.040	1.000
Tangible dummy	0.071	−0.001	1.067	0.999
Intangible dummy	0.118	0.000	2.095	0.998
Dummy export	0.232	0.000	3.981	1.002
<i>Manuf. and train:</i>				
Sales growth	−0.002	−0.036	0.268	0.868
Dummy individual	−1.107	−0.005	0.479	1.001
Tangible dummy	0.535	0.060	1.248	1.049
Intangible dummy	0.260	−0.031	3.936	0.810
Export dummy	0.367	−0.024	6.574	0.848

Notes: Panel A reports the IPWRA first-stage Multinomial Logit coefficients. Baseline omitted category is “non-manufacturing, non-training firms”. Robust standard errors in parentheses, clustered by firm. Asterisks denote significance levels (* at 10%, ** at 5%, and *** at 1%). Panel B reports raw and IPW-weighted observations by group, together with standardized differences and variance ratios for each covariate, relative to the baseline category.

- Bryan, J., 2006. Training and performance in small firms. *Int. Small Bus. J.* 24 (6), 635–660.
- Cappelletti, M., Giuffrida, L.M., 2022. Targeted Bidders in Government Tenders. ZEW Discussion Papers 22–030, ZEW - Leibniz Centre for European Economic Research.
- Caravella, S., Crespi, F., Guarascio, D., Tubiana, M., 2021. Heterogeneity in the demand-growth relationship at the firm level: the role of demand sources and innovation/knowledge characteristics. *Econ. Innov. New Technol.* 30 (5), 516–535.
- Cattaneo, M.D., 2010. Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *J. Econometrics* 155 (2), 138–154.
- Cattaneo, M.D., 2018. Multi-valued treatment effects. In: *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, London.
- Cattaneo, M.D., Drukker, D.M., Holland, A.D., 2013. Estimation of multivalued treatment effects under conditional independence. *Stata J.* 13 (3), 407–450.
- CEDEFOP, 2015. Who Trains in Small and Medium-Sized Enterprises : Characteristics, Needs and Ways of Support. Cedefop Research Paper 50, Publications Office of the EU.
- Coad, A., Tamvada, J., 2012. Firm growth and barriers to growth among small firms in India. *Small Bus. Econ.* 39 (2), 383–400.
- Corrado, C., Hulten, C., Sichel, D., 2005. Measuring capital and technology: An expanded framework. In: Corrado, C., Hulten, C., Sichel, D. (Eds.), *Measuring Capital in the New Economy*. In: *Studies in income and wealth*, vol. 65, University of Chicago Press.
- Cosh, A., Duncan, J., Hughes, A., 1998. Investment in Training and Small Firm Growth and Survival: An Empirical Analysis for the UK 1987–95. Research Report 36, UK Department for Education and Employment.

- Dearden, L., Reed, H., Van Reenen, J., 2006. The impact of training on productivity and wages: Evidence from british panel data. *Oxford Bull. Econ. Statistics* 68 (4), 397–421.
- Devins, D., Gold, J., Johnson, S., Holden, R., Matlay, H., 2005. A conceptual model of management learning in micro businesses. *Educ. + Train.* 47 (8), 540–551.
- Devins, D., Johnson, S., 2003. Training and development activities in SMEs. *Int. Small Bus. J.* 21 (2), 213–228.
- Dosi, G., 2023. Innovation as an evolutionary process. In: *The Foundations of Complex Evolving Economies: Part One: Innovation, Organization, and Industrial Dynamics*. Oxford University Press.
- Dosi, G., Grazzi, M., Tomasi, C., Zeli, A., 2012. Turbulence underneath the big calm? The micro-evidence behind Italian productivity dynamics. *Small Bus. Economics* 39 (4), 1043–1067.
- Dosi, G., Guarascio, D., Ricci, A., Virgillito, M.E., 2021. Neodualism in the Italian business firms: training, organizational capabilities, and productivity distributions. *Small Bus. Econ.* 57 (1), 167–189.
- Dosi, G., Nelson, R.R., Winter, S.G., 2000. *The Nature and Dynamics of Organizational Capabilities*. Oxford University Press.
- Farvaque, N., Voss, E., Lefebvre, M., Schutze, K., 2009. Guide for training in SMEs. In: *Rapport Final Pour la Commission Européenne (DG Emploi, Affaires Sociales et Égalité Des Chances)*, Juin.
- Gherhes, C., Williams, N., Vorley, T., Vasconcelos, A.C., 2016. Distinguishing micro-businesses from SMEs: A systematic review of growth constraints. *J. Small Bus. Enterp. Dev.* 23 (4), 939–963.
- Hall, B.H., Lotti, F., Mairesse, J., 2009. Innovation and productivity in SMEs: empirical evidence for Italy. *Small Bus. Economics* 33 (1), 13–33.
- Henley, A., Song, M., 2020. Innovation, internationalisation and the performance of microbusinesses. *Int. Small Bus. J.* 38 (4), 337–364.
- Imai, K., Ratkovic, M., 2014. Covariate balancing propensity score. *J. R. Stat. Soc. Ser. B* 76 (1), 243–263.
- Imbens, G.W., 2000. The role of the propensity score in estimating dose-response functions. *Biometrika* 87 (3), 706–710.
- Imbens, G.W., Rubin, D.B., 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge Books, Cambridge University Press, no. 9780521885881.
- Jiang, K., Takeuchi, R., Lepak, D.P., 2013. Where do We Go From Here? New Perspectives on the Black Box in Strategic Human Resource Management Research. *J. Manage. Stud.* 50 (8), 1448–1480.
- Kelliher, F., Reinl, L., 2009. A resource-based view of micro-firm management practice. *J. Small Bus. Enterp. Dev.* 16 (3), 521–532.
- Koch, M.J., McGrath, R.G., 1996. Improving labor productivity: Human resource management policies do matter. *Strateg. Manag. J.* 17 (5), 335–354.
- Konings, J., Vanormelingen, S., 2015. The impact of training on productivity and wages: Firm-level evidence. *The Rev. Econ. Stat.* 97 (2), 485–497.
- Kotey, B., Slade, P., 2005. Formal human resource management practices in small growing firms. *J. Small Bus. Manag.* 43 (1), 16–40.
- Lai, Y., Saridakis, G., Johnstone, S., 2017. Human resource practices, employee attitudes and small firm performance. *Int. Small Bus. J.* 35 (4), 470–494.
- Liberman-Yaconi, L., Hooper, T., Hutchings, K., 2010. Toward a model of understanding strategic decision-making in micro-firms: Exploring the Australian information technology sector. *J. Small Bus. Manag.* 48 (1), 70–95.
- Linden, A., Uysal, S.D., Ryan, A., Adams, J.L., 2016. Estimating causal effects for multivalued treatments: A comparison of approaches. *Stat. Med.* 35 (4), 534–552.
- Lopez, A., Neves, P., Cunha, M., 2019. A high-growth firm contingency test of the formalization–performance relationship. *J. Small Bus. Manag.* 52 (S2), 347–396.
- Mancinelli, S., Mazzanti, M., 2009. Innovation, networking and complementarity: evidence on SME performances for a local economic system in North-Eastern Italy. *The Ann. Reg. Science* 43 (3), 567–597.
- Martins, P.S., 2021. *Employee Training and Firm Performance: Evidence from ESF Grant Applications*. OECD Productivity Working Papers 23, OECD Publishing, Paris.
- McCaffrey, D.F., Ridgeway, G., Morral, A.R., 2004. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychol. Methods* 9 (4), 403–425.
- Mitri, S.D., Socio, A.D., Russo, P.F., Nigro, V., 2013. Micro-Enterprises in Italy: A First Analysis of Economic and Financial Conditions. *Questioni di Economia e Finanza (Occasional Papers)* 162, Bank of Italy, Economic Research and International Relations Area.
- Mohrenweiser, J., Zwick, T., Backes-Gellner, U., 2019. Poaching and firm-sponsored training. *Br. J. Ind. Relations* 57 (1), 143–181.
- OECD, 2017. *Entrepreneurship at a Glance 2017*. OECD Publishing, Paris.
- Palmieri, R., 2016. Sfruttamento delle fonti di dati di qualità presenti sul territorio per l'arricchimento ed il miglioramento della produzione statistica ufficiale italiana riguardante la piccola e micro impresa: il progetto TREND. Paper presented at the XXXVII Conferenza di Scienze Regionali.
- Patton, D., Marlow, S., Hannon, P., 2000. The relationship between training and small firm performance; research frameworks and lost quests. *Int. Small Bus. J.* 19 (1), 11–27.
- Reid, G.C., 1993. The state of british enterprise: Growth, innovation and competitive advantage in small and medium sized firms : Small business research centre, university of cambridge. *Int. J. Ind. Organ.* 11 (1), 147–150.
- Reid, G., 2007. *The Foundations of Small Business Enterprise: An Entrepreneurial Analysis of Small Firm Inception and Growth*. Routledge.
- Rodrigues, A.C., Carvalho, H., Caetano, A., Santos, S.C., 2022. Micro-firms way to succeed: How owners manage people. *J. Bus. Res.* 150, 237–248.
- Roth, F., Sen, A., Rammer, C., 2023. The role of intangibles in firm-level productivity – evidence from Germany. *Ind. Innov.* 30 (2), 263–285.
- Sels, L., Winne, S., Delmotte, J., Maes, J., Faems, D., Forrier, A., 2006. Linking HRM and Small Business Performance: An Examination of the Impact of HRM Intensity on the Productivity and Financial Performance of Small Businesses. *Small Bus. Econ.* 26 (1), 83–101.
- Sheehan, M., 2014. Human resource management and performance: Evidence from small and medium-sized firms. *Int. Small Bus. J.* 32 (5), 545–570.
- Słoczyński, T., Uysal, S.D., Wooldridge, J.M., 2022. Doubly robust estimation of local average treatment effects using inverse probability weighted regression adjustment.
- Storey, D., 2004. Exploring the link, among small firms, between management training and firm performance: A comparison between the UK and other OECD countries. *Int. J. Hum. Resour. Manag.* 15 (1), 112–130.
- Stuart, T.E., Yim, S., 2010. Board interlocks and the propensity to be targeted in private equity transactions. *J. Financ. Econ.* 97 (1), 174–189.
- Webster, E., Jensen, P.H., 2006. Investment in intangible capital: An enterprise perspective. *Econ. Rec.* 82 (256), 82–96.
- Wennekers, S., Thurik, R., 1999. Linking entrepreneurship and economic growth. *Small Bus. Economics* 13 (1), 27–56.
- Westhead, P., Storey, D., 1996. Management training and small firm performance: Why is the link so weak? *Int. Small Bus. J.* 14 (4), 13–24.
- Wooldridge, J.M., 2007. Inverse probability weighted estimation for general missing data problems. *J. Econometrics* 141 (2), 1281–1301.
- Wooldridge, J.M., 2010. *Econometric Analysis of Cross Section and Panel Data*, second ed. MIT Press, Cambridge, MA.
- Wright, M., Roper, S., Hart, M., Carter, S., 2015. Joining the dots: Building the evidence base for SME growth policy. *Int. Small Bus. J.* 33 (1), 3–11.
- Yazdanfar, D., 2013. Profitability determinants among micro firms: evidence from Swedish data. *Int. J. Manag. Finance* 9 (2), 151–160.
- Zwick, T., 2006. The impact of training intensity on establishment productivity. *Ind. Relat.: A J. Econ. Society* 45 (1), 26–46.