The role of occupational skill sets in the digital transformation: how it progress shapes returns to specialization and social skills

Kiener, Fabienne ; Eggenberger, Christian ; Backes-Gellner, Uschi

DOI: https://doi.org/10.1007/s11573-023-01153-9

Originally published at:
Kiener, Fabienne; Eggenberger, Christian; Backes-Gellner, Uschi (2024). The role of occupational skill sets in the digital transformation: how it progress shapes returns to specialization and social skills. Journal of Business Economics / Zeitschrift für Betriebswirtschaft, 94(1):75-111.
DOI: https://doi.org/10.1007/s11573-023-01153-9

Journal of Business Economics
https://doi.org/10.1007/s11573-023-01153-9

ORIGINAL PAPER

The role of occupational skill sets in the digital transformation: how IT progress shapes returns to specialization and social skills

Fabienne Kiener[1] · Christian Eggenberger[1] · Uschi Backes-Gellner[1]

Accepted: 15 March 2023

Abstract

Workers' occupational skill sets play a crucial role in successfully handling digital transformation. We investigate whether and how different types of occupational skill sets benefit from digital transformation. We theoretically and empirically analyze wage returns of workers in occupations with more or less specialized skill sets and with more or less social skills when IT increases in their industry. Applying natural language processing methods to the texts of occupational training curricula, we develop measures for occupational specialization and social skills. We use vocational education and training curricula from Switzerland because they cover approx. two-thirds of the working population. Using curricula, industry-level IT data and individual-level administrative wage data, our individual fixed-effects analyses show that IT progress leads to higher wage returns for workers in highly specialized occupations but not for workers in more general occupations. In addition, we find that high levels of social skills cannot make up for this difference when IT advances. However, our results *indicate* that for workers with high specialization, a combination with high social skills generates additional benefits when IT advances. Overall, our results suggest that, contrary to typical assumptions in educational policy debates, workers with specialized occupational skill sets—possibly in combination

This study is funded partly by the SNSF within the framework of the national research program "Digital Transformation" (NRP 77) and partly by the Swiss State Secretariat for Education, Research, and Innovation (SERI) through its "Leading House VPET-ECON: A Research Center on the Economics of Education, Firm Behavior and Training Policies".

This paper has benefitted from many audiences and exchanges. In particular, we would like to thank the anonymous referees for the helpful comments as well as the following individuals for their help and support: Eric Bettinger, Thomas Dohmen, Simon Janssen, Uli Kaiser, Harald Pfeifer, Jens Mohrenweiser, Dinand Webbink, conference participants at SASE, at LISER-IAB conference, at SFUVET VET Congress 2022, at PILLARS, at COPE 2022, at First International Leading House Conference on the "Economics of Vocational Education and Training", at the Annual Meeting of the German Economic Association (Verein für Socialpolitik), as well as seminar participants at the University of Zurich for helpful comments and and Natalie Reid for language consulting. We thank the Swiss Federal Statistical Office for providing the labor market data.

Extended author information available on the last page of the article

Published online: 09 May 2023

with high social skills—appear to be the ones who are particularly well prepared to cope with digital transformation.

Keywords Digitalization · IT progress · Skills · Education · Human capital

JEL Classification I26 · J24 · O33

1 Introduction

Digital transformation, i.e., the rapid advancement of IT progress and increasing use of computers in the workplace, has a significant impact on how workers with different skill sets coordinate and which types of occupational skill sets are increasingly beneficial. We argue theoretically that higher specialization may be very productive but only with effective coordination. Coordination efforts may become easier with higher IT progress and increasing use of computers in the workplace, as well as with higher levels of social skills, because both IT progress and social skills may help to reduce coordination costs.

As an example, consider two workers with highly specialized skill sets, such as one worker with only one skill A (e.g., sewing) and another worker with only one other skill B (e.g., designing apparel). Having specialized skill sets makes these workers very productive in their own tasks, but this specialization requires them to collaborate to produce a marketable product (e.g., a designer piece of haute couture). To collaborate, workers must be able to communicate and coordinate effectively. Assuming that such coordination is costly, advances in IT and the increasing use of computers in the workplace make it more efficient to communicate and coordinate. Thus workers with very specialized skill sets may become more valuable with IT progress because the increased availability of IT reduces their coordination costs. However, we also argue theoretically that the same effect should not occur for workers who do not have such specialized skill sets, i.e., for workers with general skill sets (who possess both skills A and B). As their need for coordination is much lower, they can be expected to have lower benefits from advances in IT and the increasing use of computers in the workplace. Moreover, for workers who have higher coordination needs due to their specialized skill sets, social skills (which improve communication and reduce coordination costs) may be more valuable than for those with general skill sets. However, the joint effect of increasing IT progress on workers having social skills is theoretically ambiguous because, on the one hand, IT progress could increase the value of social skills and, on the other hand, IT progress could replace the need for social skills.

In this paper, we study whether and how workers in occupations with more or less specialized skill sets and with a low or high level of social skills benefit from digital transformation, i.e., from faster IT progress and increasing use of computers at workplaces in their industry. The paper starts with theoretical considerations on the relationships between IT progress, specialization, and social skills based on a combination of two strands of economic models—i.e., Becker and Murphy (1992), who study coordination costs for workers with differing degrees of specialization, and Deming (2017), who studies the effects of social skills on coordination costs—and by adding IT progress as an additional factor. By doing so, we investigate the effect of IT progress on returns to *both* specialization and social skills, a combination that theoretical and empirical studies have not yet examined. Based on our theoretical considerations, we empirically examine how IT progress affects the wages of workers with different occupational skill sets.

Our empirical analyses from individual fixed-effects estimations lead to two main results. First, we show that workers with highly specialized skill sets have increasing wage returns when IT progress is higher—no matter whether they have high or low levels of social skills. Second, our findings indicate that workers with high social skills do not receive higher wages with increased IT progress per se. However, we find some evidence that if workers with high social skills also have highly specialized skill sets, they receive higher wages when IT advances. Thus our findings indicate that a combination of high specialization and high social skills is a particularly useful skill combination in times of digital transformation, i.e., when IT use at the workplaces in their respective industries increases.

We derive these results from empirical analyses based on a combination of three datasets: (1) We use text data from training curricula of Swiss vocational education and training (VET) programs because they allow us to extract detailed information on workers' skill sets; they provide proxies for the level of specialization and the level of social skills for all middle-skilled workers on an occupational basis. The main advantage of using training curricula texts is that they allow us to measure skill sets in great detail by applying novel natural language processing methods (NLP), i.e., we use machine-learning methods for text analyses to extract detailed occupational skill sets. Swiss VET curricula are an ideal data source for two main reasons: First, as two-thirds of the Swiss labor force acquire a VET diploma, these curricula cover basically all middle-skilled workers, i.e., a major part of the whole labor market of a highly developed and innovative country. Second, Swiss VET curricula are very detailed with an average of 44 pages, they describe all skills that workers in a particular VET occupation possess, and previous research has shown that proxying skills of middle-skilled workers by using such occupational training curricula allows predicting individual labor market outcomes (Eggenberger et al. 2018; Eggenberger and Backes-Gellner 2023; Kiener et al. 2022b, 2023). A recent working paper by Langer and Wiederhold (2023) has applied a similar method to German VET curricula and finds that proxing skills of middle skilled workers based on VET curricula also helps to explain labor market outcomes in Germany.

(2) To proxy IT progress, we use data from EUklems. Specifically, we use data on tangible stocks in computing equipment (i.e., IT capital stocks) across industries over time.[1] To avoid endogeneity problems that could arise because IT progress could depend on available skills and may thus not be fully exogenous to available skills, we use non-Swiss industry-specific IT data as a proxy for an exogenous

[1] Data stem from https://euklems.eu/download/ provided by the Vienna Institute for International Economic Studies (see methodological report by Stehrer et al. 2019).

measure of IT progress; i.e., we use data for Germany as the largest neighboring country and with the most similar industry structure to Switzerland.

(3) We use administrative labor market data (the Swiss Social Protection and Labor Market data, SESAM[2]) to measure the outcomes of an interplay between IT progress and different occupational skill sets at an individual worker level.

Our paper makes two main contributions. First, we add novel empirical findings on returns to occupational skill sets with high vs. low levels of specialization and high vs. low levels of social skills during times of IT progress. Such findings have been very limited so far.[3] Second, we provide evidence of the interplay of specialization and social skills in times of IT progress. While previous research (Deming 2017; Deming and Kahn 2017; Weinberger 2014) and a later paper by Langer and Wiederhold (2023) also show that returns to social skills have increased in the long run, there is, so far, no evidence of how these returns are affected by differing degrees of IT progress—and particularly not on how a combination of social skills with a more or less specialized skill set is affected. In addition, Langer and Wiederhold (2023) examine in their recent working paper—applying a similar method to German curricula for measuring skills—that, besides returns to social skills, returns to digital skills also have increased over time.While their paper provides insights on long-term returns to digital skills, it does not examine how IT progress in the respective industry affects returns to skill sets.[4] Our paper shows how workers with different types of skill sets (social skills and a more or less specialized skill bundle) are affected by changes in IT progress, i.e., in the context of digital transformation.

Our findings also lead to important policy implications for firms, educational policymakers, and training institutions such as vocational schools, technical colleges or even universities, which have to develop future-oriented professional training curricula. While so far, a piece of typical policy advice in the context of IT progress was to make skill sets more general rather than more specialized in order to increase the adaptability of workers (see e.g.,Green 2002; Jansen et al. 2017), our findings *indicate* that—when IT advances, thereby decreasing coordination costs—high levels of specialization are not necessarily a disadvantage, likely even more so if they are combined with higher levels of social skills.

2 Theoretical considerations and hypotheses on the relationship between IT progress, specialization and social skills

To theoretically understand how IT progress may affect workers with different types of skill sets, we combine two strands of economic models of production (Becker and Murphy 1992; Deming 2017) and add IT progress as an additional factor. From

[2] SESAM, which follows a rolling panel structure, covers a representative sample of the population in Switzerland and comprises administrative wage data and information about individuals' sociodemographic characteristics and education.

[3] An exception is, for example, Borghans and ter Weel (2006), who focus on specialization within firms but not on workers with specialized skill sets across the entire labor market.

[4] In contrast, Langer and Wiederhold (2023) use regional skill demands from job ad data in addition to curricula and work with the assumption that these skill demands can also proxy firm technology.

Becker and Murphy (1992), who study specialization, productivity and the costs of coordination for differing degrees of specialization, we use the insight that increased specialization may ceteris paribus increase productivity while also increasing coordination costs. From Deming (2017), who studies the effects of social skills on coordination costs, we use the insight that social skills may decrease coordination costs because better communication facilitates collaboration. We add IT progress as another factor that supports communication, thereby decreasing coordination costs.[5] The following illustrative example explains how these three factors influence and complement each other (for the mathematical modeling of the relationships, see Appendix 1).

Consider a production that has one output (a jacket) and requires two skills (design and tailoring of the jacket). We assume a worker *A* with one production skill "design" is more productive in this skill than an otherwise identical worker *Z* with two production skills "design" and "tailoring." The same applies to a worker *B* with the one production skill "tailoring." Thus team production of two fully specialized workers *A* and *B* could be more efficient than the production of two workers *Z*, but only if the two specialized workers *A* and *B* coordinate their outputs across the different production steps. When the two fully specialized workers *A* and *B* combine their outputs, coordination costs occur. Now consider workers specialize even more: the tailoring task of worker *B* is now further divided across two different workers *C* and *D*, one who only cuts fabrics and one who only sews. Workers *C* and *D* are more productive in their respective tasks than worker *B*, but the overall coordination costs increase even more. The group of workers (*A*, *C*, and *D*) now all have to communicate to transform a design into a fitted cut and a sewn-together jacket. If workers become even more specialized, their individual productivity increases further but so do coordination costs. Thus there is a trade-off between specialization and coordination costs (for a formal derivation of this proposition, cf., Becker and Murphy 1992, and our model in Appendix 1). This trade-off determines the optimal level of specialization in a worker's skill set.

Now consider the introduction of different levels of social skills. Social skills determine the ability to interact well with others; they improve collaboration. Thus, if workers possess higher levels of social skills, these skills help to decrease coordination costs (as shown in Deming 2017). Consequently, introducing social skills changes the optimal skill set (as shown in our model in Appendix 1, combining Deming 2017 and Becker and Murphy 1992).

In addition, now also consider *IT progress,* e.g., an introduction of computer equipment, emailing or 3D design software that makes communication and exchange of data and plans (design outlines, cutting patterns, sewing plans, production scheduling, and so forth) easier. Such IT equipment also decreases coordination

[5] Our model differs from that of Borghans and ter Weel (2006), who formulate a model that combines IT adoption in firms and specialization but does not include social skills. Moreover, our theoretical premises differ from those in the organizational literature, e.g., Caroli and van Reenen (2001); Dessein and Santos 2006); Lindbeck and Snower (2000), which focusses on firms' organizational changes in response to digital transformation.

costs. Thus increasing the availability of IT equipment also changes the optimal skill set.

Based on these theoretical considerations, we can derive one hypothesis on the relationship between IT progress and workers' specialization on productivity and two contrasting hypotheses on the relationship between IT progress and social skills on productivity. The hypotheses are:

(H1) Increasing IT equipment in the workplace makes the coordination between specialized workers easier and therefore makes a higher specialization beneficial to workers.
(H2a) Increasing IT equipment makes the social skills of workers more valuable (because with increasing IT, a higher specialization of workers is more efficient and—as a complement—workers also require more social skills to better communicate = scenario 1).
(H2b) In contrast, increasing IT equipment could also substitute social skills—instead of complementing them—and in this case, would not increase the value of social skills (because communication via IT replaces the need for workers' social skills = scenario 2).

Which of the two scenarios is relevant in the real world is a priori unclear and remains an empirical question. Thus we have three hypotheses to test in our empirical part.

To operationalize the skill sets of workers in our empirical analyses, we assume that the *skill sets* of workers are given by their occupation. Workers in each occupation (e.g., the "designer" or the "tailor") possess a distinct skill set that is defined by the training curriculum of the respective VET occupation. In countries with occupational labor markets (such as Switzerland and other countries with VET), it can be further assumed that workers transfer their occupational skill sets across the firms they work for (as pointed out e.g., by Marsden 1999). Thus assuming that the skill sets of workers are given by the occupation of the worker is a particularly reasonable assumption in an empirical setting like Switzerland, which we investigate in this paper (see Sect. 3 explaining the setting and datasets). Moreover, in our empirical application, we make the simplifying assumption that all workers within an occupation have an identical degree of specialization or of social skills, i.e., the levels of specialization and the levels of social skills vary across occupations, but they are identical within the same occupation. Based on these assumptions, we can investigate the empirical question of how IT progress affects returns to more or less specialized occupations and with higher or lower levels of social skills.

The following Sect. 3 introduces our datasets, and Sect. 4 discusses how we apply NLP methods to extract the skill sets, which we need for our specialization measure, from the VET training curricula.

3 Data

For our empirical analyses, we need data to measure the specialization of skill sets, the level of social skills in occupations and annual industry-specific IT progress, as well as individual workers' labor market data. We use three different datasets: 1. occupational training curricula to measure occupational specialization and social skills, 2. data on the growth of IT capital stock to measure IT progress, and 3. labor market data for workers' individual labor market outcomes and individual controls. We use Switzerland as our empirical setting. The Swiss setting has two main advantages: First, for Switzerland, we have very detailed upper-secondary VET curricula information to reliably extract skill sets. Second, in Switzerland, the labor market for workers at a middle-skilled level is more or less completely defined by VET graduates, given that two-thirds of Swiss workers graduate from an upper-secondary VET program and thus acquire the occupational skill sets we derive from the curriculum texts. Moreover, previous studies have used similar approaches for analyzing skills described in Swiss VET curricula and showed their importance for labor market outcomes (Eggenberger et al. 2018; Eggenberger and Backes-Gellner 2023; Kiener et al. 2022b, 2023). Similar features are found in a later study on German curricula by Langer and Wiederhold (2023).

The subsequent paragraphs describe each dataset and discuss why they are suitable for our empirical analyses. For data on specialization and social skills, we use occupational skill sets from Swiss upper-secondary VET programs. These VET programs last three to four years and they combine on-the-job training with school-based training. Their detailed curriculum texts describe in great detail all the skills that each occupation requires (per curriculum, the text has on average approx. 44 pages). The curricula list the legally binding learning goals for the skills that the VET apprentice must learn, and various examinations guarantee that apprentices receiving a VET diploma have indeed acquired those prescribed skills.

We use the curriculum texts to measure, first, occupational specialization, which is defined by the extent to which the production skills in an occupation are very focused on only a few selected tasks (rather than on many different tasks) and, second, the level of social skills in an occupational skill set. Curriculum texts describe both production and social skills, but they do so by using different semantic concepts. A curriculum consists of many learning goals, which describe the production skills an apprentice has to acquire. While one learning goal usually describes only one production skill, such a description can nevertheless also describe a social skill simultaneously. For example, consider the following learning goal: "The apprentice advises customers in a friendly manner." This learning goal, on the one hand, describes the production skill "customer service" and, on the other hand, this text is also about friendliness and implies that simultaneously a social skill is taught. Given that production skills and social skills occur in semantically different ways, we use two different machine-learning methods to measure the different types of skills in the curricula texts.

To extract the production skills needed for our specialization measure, we apply novel NLP methods to our curriculum text database. Using these text data, we construct our occupational skill measures in line with Eggenberger et al. (2018), who manually derived skills data sets from the learning goals of the curriculum texts. Our curriculum database contains, on average, approx. 134 learning goals per occupation—in total 21,776 learning goals in 163 occupations.[6] Section 4 shows how we extract production skills from each curriculum. Section 5.1 afterwards describes how we construct the specialization measure based on the skills retrieved in the previous step.

To identify social skills in curricula, we use the social skill measures developed by Kiener et al. (2023), who already used NLP methods to detect social skills in text passages of VET curricula. Typically in these curricula, the text passages on social skills include communication skills, the ability to work in a team, or being friendly. For example, in the curriculum of a bookseller, the algorithm identified social skills in the following text passage (in italics): "Booksellers attach importance to … *successful communication.*" Thus this step allows us to extract the amount of text passages that describe social skills in each curriculum. Based on this information, we build the relative proportion of the text passages that describe social skills in comparison to text passages that describe the other skills from the previous step. The proportion is used to define the level of social skills in each curriculum.

For data on IT progress, we follow the economics literature that uses firms' investment in IT capital to proxy IT progress (e.g., Gaggl and Wright 2017).[7] For example, previous macro-economic studies use national accounts data on IT capital investments to analyze how productivity growth is driven by IT (e.g., Gordon and Sayed 2019). In line with the reasoning that firms' investment in IT capital proxies IT progress, we use the EUklems dataset, which contains annual information on tangible stocks in computing equipment (i.e., IT capital stocks) from national accounts data for the period 1995–2017. In particular, we use the data for "computing equipment" from the EUklems data because it matches most closely with our theoretical reasoning. Our theoretical reasoning draws on coordination that primarily depends on the availability of new computer equipment at the workplace. We use data at the industry level because such yearly data is not available at the firm level; thus we assume that workers in an industry in which firms overall invest a lot in IT capital are also, on average, subject to large IT progress at their particular firm. To measure IT progress in different industries across time, we use annual changes in IT capital stocks (see Sect. 5.1 for more details on the measure). The main advantage of this data is that it is national account data provided by the national European statistical institutes and is, therefore, highly reliable.

For Switzerland, such national accounts data for IT capital stocks (or any similar data) is not available across industries over time.[8] Therefore, we use the best

[6] More precisely, the number of average learning goals are 133.595 (thus $133.595 \times 163 = 21{,}776$).

[7] In a regression discontinuity design, Gaggl and Wrigth (2017) investigate the effect of a tax reduction on firms' IT investments. However, their focus is on non-routine, cognitive-intensive workers and not on specialization and social skills.

[8] The Swiss Federal Statistical Office does not collect such data at the industry level. Other data sources, such as firm surveys, which contain industry information, are not available at the annual level over two decades.

alternative, data from Germany. Given that we use a growth variable of IT capital stock, different industry sizes between Germany and Switzerland do not play a role in our IT progress measure. The only assumption we need to make is that the growth of IT capital within a given industry is similar in Germany and Switzerland, which seems likely given that they both compete in the same markets and are similarly developed and innovative economies. In addition, using German data has one big advantage because it ensures that the IT proxies that we use are exogenous to the skills structure in Switzerland as explained in more detail in Sect. 5.3.2 below.

To study the effect of IT progress on returns for workers, we also need reliable outcome data. Thus, as a third dataset, we use administrative labor market data that also include some individual characteristics of the workers that we can use as additional controls in our regressions. We use SESAM, which links individual survey data from the Swiss labor force survey (SLFS) with individual administrative data (e.g., wages). SESAM comprises a representative sample of the Swiss population and has a rolling panel structure. From 1999 to 2010, individuals are followed over five consecutive years. After 2010, the survey interval was shortened, following individuals for only two consecutive years. We use all available years in SESAM, i.e., from 1999 onwards until 2016 (the last available year to calculate our IT progress variable from EUklems), to benefit from as many observations of individuals as possible. SESAM provides us with all information on annual wage data, educational qualification, current occupation, and current industry. For our estimations, we use a sample of individuals with a VET diploma as their highest education. We use the current occupation of these workers to assign their occupational skill set, i.e., we link each worker with the occupational specialization and social skills that we derive from the respective curriculum texts. We chose the current occupation (rather than the originally learned occupation) for two main reasons: First, this approach is consistent with previous papers using the skills of the current (and not the trained) occupation to identify the skills that workers have when working in their current workplace (cf., Rinawi and Backes-Gellner 2021 and Eggenberger et al. 2022). Second, according to our theoretical reasoning, we are interested in the types of skills that workers have when working in their current workplace; these skills of the occupation they work in can be more or less specialized and can, therefore, be expected to be more or less affected by IT progress. As we are not studying skills from a training perspective but from a workplace perspective, i.e., we are interested in the skill sets that workers typically have while working in a particular workplace and how these are affected by more or less exposure to IT progress, we decided to also use the current occupation and to take the typical skill sets of that occupation to approximate the workers' skills in their current workplace. By doing so, we assume that firms assign workers only to workplaces for which workers fulfill the respective skills requirements, i.e., they have the typical occupational skills as defined by the training curricula of the respective occupations, which is consistent with previous literature as mentioned above.

In our final dataset, we are able to link 19 industries and 111 occupations wherever a matching between the SESAM data and the EUklems data, and between the

SESAM data and the curricula, is possible.[9] Our final dataset has a rolling panel structure and covers the years 1999 (the first available year in SESAM) through 2016 (the last available year to calculate our IT progress variable from EUklems). Before we discuss our empirical framework, we describe how we extract production skills to measure specialization in the curricula because our methodological approach is novel and has not yet been applied to curricula.

4 Measuring production skills in curricula

This section discusses how we use NLP methods to extract production skills from Swiss VET curricula[10]—which we need to construct our specialization measure in Sect. 5.1. To extract the production skills, we start by copying the plain texts of the learning goals of all curricula into a text database (see Eggenberger et al. 2018 for more details on the structure of the curricula).[11] For our 163 occupations, this procedure leads to a database of 21,776 learning goals. These learning goals will be the unit of analysis in our skill extraction procedure (thus learning goals are our "documents").[12]

The main goal of our skill extraction procedure is to identify an exhaustive set of all production skills that appear in any training curriculum and calculate the weight of each skill per curriculum. Therefore, we want to identify *sets* of learning goals, i.e., skill categories, that are similar, both within (i.e., a curriculum comprises learning goals that are encompassed in the same skill category) and across (i.e., different curricula comprise learning goals that are encompassed in the same skill category) training curricula. To achieve this goal, we proceed in three steps. In the first step, we transfer each learning goal into a multidimensional vector representation or document "embedding." Document embeddings are an NLP method to encode semantic meaning in mathematical form. In the second step, these embeddings allow us to cluster the learning goals, creating groups of similar learning goals. We

[9] We do an exact matching of SESAM occupations with the occupations from the curricula texts, which in some cases have more detailed occupational fields. Thus when matching the curricula occupations to the SESAM classification we lose some occupational details, which result in fewer occupations in the final data set than the number of occupations in our curriculum database.

[10] The NLP procedure used in this paper is similar to the one in Eggenberger and Backes-Gellner (2023) but specifically tailored to measure specialization.

[11] Following Eggenberger et al. (2018), we take all third-level learning goals in the curricula (see Eggenberger et al. (2018), for the detailed explanation of these learning goals). Moreover, we remove all words or phrases that contain occupation titles from the database. This removal ensures that the algorithms in the next steps will treat each learning goal as independent, even if learning goals originate from the same occupational curriculum.

[12] Splitting up the curriculum texts into the lowest level learning goals and using these texts as the unit of analysis has the advantage that we deal with texts that are much more coherent than the texts we would get when analyzing the curriculum texts on an aggregated level. A drawback of this approach is that data sparsity under the short-text scenario hinders the process of finding document-topic distributions due to the lack of word co-occurrence information (Yi et al. 2020). However, as we will discuss in this section, this problem can be overcome by using large external databases to gather information about the internal semantic relationships of the words and leverage this information in our analysis.

will interpret each of these clusters as a distinct "skill category." In the third step, we calculate the weight of each skill category in each curriculum.

To create the document embeddings in the first step, i.e., a vector representation of each learning goal, we use a sentence transformer network model. Transformer models leverage large, unlabeled repositories of texts (such as Wikipedia or books) to create meaningful sentence embeddings such that sentences with similar meanings are close in vector space (Reimers and Gurevych 2019). More precisely, we use a transformer model that is based on Bidirectional Encoder Representations from Transformers (BERT)[13] (Devlin et al. 2018). BERT is a bidirectionally trained language model developed by researchers at Google AI language. In contrast to previous embedding methods (such as word2vec), BERT is fully context-dependent, meaning that words (or sentences) will have different vector representations depending on the context in which they appear. We average all sentence vectors for a given learning goal to obtain one vector per learning goal.[14]

In the next step, we proceed to cluster similar learning goals. For the clustering, we use the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm. The algorithm aims at finding clusters of vectors (in our case, each learning goal is a vector) in a high-dimensional vector space. Being a density-based clustering approach, HDBSCAN looks for regions in the data that are denser (i.e., linguistically more coherent) than the surrounding space. This approach has the advantage that we do not have to make any implicit assumptions about the shape of the clusters we are looking for. HDBSCAN thus works particularly well for clusters of arbitrary shapes with different sizes and densities (Campello et al. 2013) and outperforms classical clustering algorithms in many real-life applications, including text clustering. Another important advantage of HDBSCAN is that it does not require specifying the number of clusters we are looking for in advance; instead, we merely have to define a minimum cluster size.[15] We set the minimum cluster size to 15, meaning that the algorithm has to find at least 15 closely related learning goals in order to form a cluster, i.e., a skill category. We chose this minimum cluster size to have a reasonably defined size of skill categories. As the average curriculum has about 130 learning goals, a skill category that occurs in one curriculum only would have to contain more than 10% of this curriculum's learning goals to be considered an independent skill. This procedure leads to the identification of 190 well-identified clusters, i.e., skill categories, in our data, which is as expected when

[13] BERT has achieved state-of-the-art results in a wide variety of NLP tasks, including sentence similarity (Reimers and Gurevych 2019). We use a freely available, multiple-language BERT model (python package *BERTopic*), which was optimized for semantic textual similarity tasks. We encode each sentence of each learning goal as a vector. The BERT model uses 768 dimensions for one vector.

[14] The procedure and reason for averaging sentence vector are as follows: Most learning goal in Swiss VET curricula consists of one sentence, but sometimes a learning goal consists of a few sentences that describe a similar topic and have a similar meaning. If a learning goal consists of two sentences, each of these sentences is represented by a mathematical vector. Both vectors are used to define the topic of such a learning goal, i.e., we take the average of the two mathematical vectors, thus generating the most general representation of the whole learning goal, which corresponds to the standard procedure in NLP text analyses.

[15] In principle, HDBSCAN has a few additional parameters to specify; however, these rarely have a practical effect on the clustering outcome.

comparing this number with the previous manual analysis of older curricula from Eggenberger et al. (2018), who manually identified 181 skills.

In the third and last step, we calculate the relative importance of each skill category per curriculum. We weigh the learning goals in a curriculum that are assigned to one particular skill category with the inverse of the total number of learning goals in the curriculum, calculating the share of each skill category per curriculum. The resulting skill database will serve as the input data to calculate the occupational specialization measure (see Sect. 5.1).

To illustrate the content of the skill categories, we apply an automated procedure to extract representative keywords for each category.[16] While the content of the skill categories (i.e., the keywords) is not needed for our measure of specialization, these descriptions are nevertheless helpful to illustrate how the algorithm works. Table 1 lists some of the generated skill categories and the corresponding representative keywords. Each skill category receives a number that is an identifier. Besides the skill number, the table lists the most important keywords (i.e., skill descriptions), the frequency (i.e., in how many curricula the skill category occurs), the mean and maximum weight (i.e., average and maximum weight of the skill category across curricula, in which the skill category occurs in). For example, skill number 37 in the table is a skill category on nutrition, as the keywords indicate. This skill category occurs in 16 curricula and has a mean weight across those curricula of 0.045. This means that the nutrition skill category is only relevant in a few curricula—compared to the other examples—but if it occurs in a curriculum, it is an important skill category. Another example of a skill category is number 74 including technical drawing, which occurs in many curricula (53), but its weight is rather low on average (0.029). However, this skill category has a high maximum weight of 0.141, meaning that in at least one curriculum, this skill category is very important. These examples of skill categories emphasize that the algorithm leads to reasonable and well-interpretable results.

The applied procedure to extract the production skills results in the weight of each skill category for each curriculum to calculate our measure of specialization (see Sect. 5.1).

[16] From each skill category, which is a clustering of a set of learning goals, we aim at displaying the description and content of these skill categories. To get the skills descriptions (i.e., keywords), we aggregate all learning goals of a given cluster (i.e., skill category) into a newly created single document. So for each skill category, we now have a document. Then we apply a term frequency-inverse document frequency (TF-IDF) analysis. TF-IDF is a numerical statistic that reflects the importance of a word in a document. The TF-IDF statistic is calculated by multiplying two metrics: the term frequency of a word in a document (how many times a word appears in a document), and the inverse document frequency of the word across all documents (how common or rare a word is in the entire document set). If a word appears in a document many times while simultaneously not appearing in any other documents, the word will get a high TF-IDF score. This procedure provides us with a set of words that *describe* a set of learning goals, i.e., a skill category. To improve the interpretability even further, we employ an additional step and eliminate words that exhibit a low coherence to the other words in the topic and overfit to the specific documents that make up the cluster. For example, some learning goals might be taught in similar locations (school, intra-firm learning centers). The name of these locations might often appear in learning goals but do in fact not contribute to the topic and are not informative for our skill categories. The last step eliminates these words using the Maximal Marginal Relevance measure (Carbonell and Goldstein 1998) and the BERT embeddings from step two.

Table 1 Illustrative examples of production skill categories extracted by NLP methods

Skill No	First keywords in English (in German) [tf-idf score]	Second keywords in English (in German) [tf-idf score]	Third keywords in English (in German) [tf-idf score]	Frequency	Mean weight	Max weight
3	Documenting quality (qualität dokumentieren) [0.46]	Test quality (prüfen qualität) [0.43]	Document measuring (dokumentieren messen) [0.42]	6	0.026	0.088
6	Comply with hygiene standards (hygienestandards einhalten) [0.70]	Hygiene hygienic standards (hygiene hygienestandards)[a] [0.70]	Hygiene standards (hygienestandards) [0.68]	3	0.027	0.061
22	Technical arithmetic (fachrechnen) [0.06]	Mathematics (mathematik) [0.05]	Calculate (berechnen) [0.04]	70	0.021	0.069
26	Polymer processing (kunststoffverarbeitung) [0.07]	Polymers apprentices (kunststoffe lernende) [0.06]	Materials polymers (werkstoffe kunststoffe) [0.05]	4	0.049	0.163
37	Nutrition (ernährung) [0.09]	Dietary formulas (kostformen) [0.03]	Nutrition principles (ernährungsgrundsätze) [0.02]	16	0.045	0.145
49	Control technology (steuerungstechnik) [0.08]	Automation (automation) [0.08]	Control SPS (steuerungen sps) [0.07]	10	0.032	0.126
74	Drawing basics (zeichnungsgrundlagen) [0.06]	Drawing technique (zeichnungstechnik) [0.06]	Drawing technique drawing basics (zeichnungstechnik zeichnungsgrundlagen) [0.04]	53	0,029	0,141
77	Safe workplace (sicher arbeitsplatz) [0.16]	Workplace design (arbeitsplatzgestaltung) [0.15]	Workplace design job (arbeitsplatzgestaltung stelle) [0.14]	24	0,023	0,067

The table shows the three keywords, uni- or bi-grams, in English—author's translation—with the highest TF-IDF score for a random sample of skill categories. In addition to the three keywords in English, the table shows in parentheses, the keywords in German and in square brackets, the TF-IDF score. The columns "Frequency," "Mean Weight," and "Max Weight" shows the distribution of the skill category across the curriculum database. The column "Frequency" reports the number of training curricula that have at least one learning goal associated with the respective skill category. Column "Mean" reports the mean weight of the skill category in curricula that have at least one learning goal associated with the respective skill category. Column "Max" reports the maximum weight of the skill category in any curriculum.

[a]The first word here in German is "Hygiene" that is translate in English to "hygiene". The second word in German ("Hygienestandards") is two words in the English translation ("hygienic standards").

5 The empirical model and methodology

After describing the data and how we extract production skills in curricula, we turn to our empirical model and operationalizations. Empirically, we analyze how—in years in which IT progress is higher—wage returns (our outcome variable) develop for workers with high specialization compared to workers with low specialization and/or for workers with high social skills—compared to workers with low social skills. To do so, we first introduce and discuss our main explanatory variables (IT progress, specialization, and social skills), discuss our sample selection, and present our empirical strategy. Subsequently, in the next section, we present our empirical results.

5.1 Explanatory variables: IT progress, specialization, and social skills

Before describing our estimation strategy, we elaborate on how we construct our main explanatory variables: IT progress, specialization, and social skills. First, the variable for IT progress is the growth of IT capital stocks in the industry the individual works. We define our variable for IT progress as IT capital stocks next year over IT capital stocks this year. The variable for IT progress varies over time across 19 industries. Our empirical strategy (individual fixed-effects model, explained in Sect. 5.3) requires that our variable for IT progress has a high variation over time and across industries, which is sufficiently fulfilled as shown in Appendix 2 (see Fig. 3 demonstrating the IT progress in every industry over time).

Second, the variable for specialization—according to our theoretical reasoning—takes a high value when an occupation focuses on only a few production skills with relatively high weight. In contrast, the variable specialization takes a low value if an occupation consists of many different skills with relatively low weights each. To condense the different weights of all skills into one single measure of specialization for each occupation, we use a Herfindahl index that is calculated as follows:

$$H_j^{190} = \sum_{k=1}^{190} \left(\frac{\text{weight of production skill } k \text{ in occupation } j}{\text{sum of all production skills in occupation } j} \right)^2 \tag{1}$$

The Herfindahl index takes the squared sum of the relative weight of each production skill. The total number of production skills that we observe across all occupations is 190 (as described in the previous section). The Herfindahl index various empirically between 0.03 and 0.25. The higher the Herfindahl index, the more specialized the occupation.

Third, the variable for social skills draws on curriculum data as the one for specialization, but the measurement is differently constructed because social skills may occur in many learning goals and across varies production skills (for more details on how social skills are measured, see Sect. 3). Thus we take the share of text describing social skills across all learning goals within a curriculum and calculate the percentage of a curriculum text that describes social skills (as in Kiener et al. 2023).

From the data of our specialization and social skills measures, we build four different types of skill sets: occupational skill sets with a high or low specialization level and occupational skill sets with a high or low level of social skills. We do so for the following reasons: First, based on our theoretical reasoning, we are interested in how IT progress affects the wage returns of workers with high specialization and of workers with high social skills. Second, we do not necessarily expect a linear relation between specialization degrees and wage outcomes or between social skill levels and wage outcomes. At the same time, we do not have the statistical power to use a large number of categories, such as percentiles, to account for nonlinearities. Taking together our theoretical interest and the empirical restrictions of our data, we decided to distinguish two groups only in each of the explanatory variables. Specifically, we construct our skill groups as follows: First, we take the sample that we use in our main regression (i.e., we only use individuals below age 45). We use each individual in this sample once for the standardization of the next step (to avoid biases due to the SESAM data structure, which would otherwise result in different numbers of observations per individual). Second, we standardize the values for the degree of specialization and the level of social skills on this sample. Third, we build two levels for each the specialization and the social skill proxy: a "high" level means the standardized value of the specialization (social skills) proxy of an occupation is above or equal to 0; a "low" level means the standardized value of the specialization (social skill) proxy is below 0. Fourth, we match these high or low levels to the occupations of the individuals.

To summarize, we construct our main explanatory variables such that the results of the estimation model show whether and how IT progress affects wage returns for workers in occupations with high or low degrees of specialization and high or low levels of social skills.

5.2 Sample and descriptive statistics

Our main sample for the empirical analysis is the younger half of the workforce because—on average—IT progress primarily affects these workers (see, for example, Bertschek and Meyer 2009 or Aubert et al. 2006, who analyze these age effects) and because we are interested in what happens when workers are affected by IT progress (depending on workers' specialization and social skills). Thus we restrict our main sample to individuals aged 19–44 with a VET diploma as their highest education level (nonetheless, in additional analyses, we also use the other age group, cf., last paragraph of result section and Appendix 2).

The summary statistics in Table 2 (see below) show descriptive statistics for this sample, i.e., the mean, the minimum, and the maximum value of annual wages, age, IT progress, high specialization, and high social skills for workers aged 19–44. Given that we have a panel dataset structure, we are interested not only in the variation of the variables across all observations (c.f., column "standard deviation overall") but also in the variation within individuals (c.f., column "standard deviation within"). Of course, the variation over all observations is higher than the one within the individual: For example, IT progress has a standard deviation of 0.1 over all observations, and within individuals, this standard deviation is lower (0.07).

Table 2 shows that in our sample, annual wages are, on average, CHF 54,887, which seems reasonable as these wages are not calculated in full-time equivalents because we want to include lower wages arising due to, for example, part-time work. The summary statistics display the minimum and maximum age in our sample, which corresponds with our age restrictions, and the average age in our sample is 33 years old. IT progress is at the mean slightly lower than 1, i.e., 0.98, meaning that on average, IT investments slightly decrease in our time span. The variables "high specialization" and "high social skills" are standardized and, therefore, take value 0 as a minimum and value 1 as a maximum, but as the standardization is done on a sample that only took one observation per individual (see Sect. 5.1), the means of the variables in the summary statistics are not exactly 0.5.

In Appendix 2, we also show the summary statistics for the same variables but by groups with high or low specialization and high or low social skills (see Table 5).[17] Moreover, correlations between the main explanatory variables are only small, as shown in Table 4 in the appendix, thereby making multicollinearity problems unlikely in our main estimations.

5.3 Empirical strategy

This section first presents the estimation model and then second elaborates on its advantages and limitations.

5.3.1 Estimation model

The estimation model analyzes wage returns as the outcome variable, regressed on the main explanatory variables, i.e., IT progress, specialization, and social skills. Specifically, the individual fixed-effects estimation model analyzes how IT progress affects wage returns for workers with high or low specialization and with high or low social skills. The estimation model thus focuses on how the wage returns of these workers change during IT progress, depending on their specialization and social skills. The estimation model is as follows:

$$\begin{aligned}\log\left(wage_{i,t}\right) = {} & a_i + \beta_1 ITProgress_{j,t} + \beta_2 ITProgress_{j,t} * HighSpecialization_o \\ & + \beta_3 ITProgress_{j,t} * HighSocialSkills_o \\ & + \beta_4 ITProgress_{j,t} * HighSpecialization_o * HighSocialSkills_o \\ & + \beta_5 year_t + \beta_6 age_{i,t} + \beta_7 age^2_{i,t} + u_{i,t}\end{aligned} \tag{2}$$

[17] Figure 3 in the appendix shows that the variation in IT progress was typically higher in the earlier years than in the later years of our time period (i.e., before and after 2009). We account for differences across years by controlling for the years in our estimations.

The outcome variable $log(wage_{i,t})$ is the log of annual wages in year (t) for individual (i).[18] The main explanatory variable, $ITprogress_{j,t}$ depends on the industry (j), in which the individual works, which may vary across time t. Thus we assume IT progress varies across industries, and individuals working in the same industry have the same (annual) IT progress measure. The other main explanatory variables, $HighSpecialization_o$ and $HighSocialSkills_o$, are measured at the occupational level of the individual and are defined by the occupation an individual works in.[19] Thus individuals working in the same occupation have the same measures for specialization and social skills.

Individual fixed effects are denoted in a_i, incorporating the influences of time-invariant differences across workers in unobserved and observed characteristics (e.g., socioeconomic status).[20] Control variables are the years of the observation, 1999 (i.e., the first year of the SESAM data) through 2016 (i.e., the last year of the IT progress variable). We also control for age and age squared.[21] Standard errors are clustered at the occupational level, the same level as the measures of specialization and social skills.[22] β_1 captures the relationship between IT progress and wage returns for workers in occupations with low specialization and with low social skills (baseline). The coefficients of the interaction terms β_2, β_3, and β_4 provide our main results because they show changes in wage returns for different groups of workers when IT advances, accounting for individual fixed effects. In our estimation model, the main variation e.g., in the interaction of IT progress and high specialization stems from changes in IT progress across time. β_2 shows the relationship between IT progress and wage returns for workers in occupations with high specialization and with low social skills, β_3 the one for workers in occupations with low specialization and with high social skills. β_4 captures the relationship between IT progress and wage returns for workers in occupations with high specialization and with high social skills. In such an interacted model, the variation between occupations and industries is important (i.e., an occupation has to occur in a number of industries). In the bar chart in Appendix 2 (see Fig. 2), we show that each occupation occurs in several industries.

[18] In contrast to Kiener et al. (2022b, 2023), we do not calculate full-time equivalents, because we compare changes of wages within an individual and in such an analysis, we consider it important to account for lower wages driven for example due to switching to part-time work.

[19] We take the individual's first occupation observed in the panel dataset for connecting the occupational skill measures to define the individual worker's skill set. We elaborate on this choice of using current occupations to match the skills to the individual labor market data in the data section.

[20] Results are robust if we include marital status as an additional, time-variant control variable in our estimations. To avoid problems of bad controls, we do however not use changes in marital status as a control variable in the main estimations.

[21] As returns to social skills or to specialization may vary across workers' age, we include age and age squared as controls variables. Moreover, our sample only includes workers below age 45.

[22] We cluster the standard errors at the occupational level because our main analysis is on groups of workers and these groups are defined by skill sets at the occupational level. This is in line with previous literature such as Eggenberger et al. (2018) or Eggenberger and Backes-Gellner (2023). Clustering at the occupational level instead of the industry level also avoids problems of having too few clusters (as for example argued in e.g., Angrist and Pischke 2009 and Cameron and Miller 2015). We avoid this problem by clustering at the occupational level.

Table 2 Summary statistics of wages, age, IT progress, specialization, and social skills

Variable	Mean	St. dev. Overall	St. dev. Within	Min	Max	N
Annual wages	54,887.36	28,867.06	7,367.82	50	984,700	55,989
Age	33.41	7.08	0.78	19	44	55,989
IT progress	0.98	0.10	0.07	0.66	1.30	55,989
High specialization	0.54	0.50	0	0	1	55,989
High social skills	0.52	0.50	0	0	1	55,989

Authors' calculations of the summary statistics individuals from the sample, data based on their skills measures, EUklems data and the SESAM, 1999–2016, for the sample of workers below age 45. High specialization is constructed after standardizing the specialization measure and taking the values equal and above zero. The same procedure is applied for the construction of high social skills

5.3.2 Advantages and limitations of our empirical strategy

Our empirical strategy has two main advantages: individual fixed effects and the exogeneity of our IT progress measure, which are also associated with some limitations. First, the individual fixed-effects estimation strategy focuses on changes in wage returns *within individuals*.[23] Thus the estimation strategy largely accounts for occupational selection concerns (e.g., different types of workers select into specialized occupations with low social skills than into less specialized occupations with high social skills). We assume that some types of workers—such as more able workers—cannot better predict in which particular year and industry IT progress will occur. If this assumption would not be valid, the interaction term of our empirical estimation could capture some selection effects. However, we argue that our assumption (that workers cannot fully foresee IT progress across industries) is highly reasonable because our measure of IT progress varies strongly over time and across industries, which is hardly foreseeable and even harder to constantly adjust to by individual workers switching industries or occupations. In sum, we argue that our individual fixed-effects approach is the main advantage of our estimation strategy.

Second, our empirical strategy requires an exogenous IT progress measure for causal inference. We argue that the assumption of the exogeneity of our IT progress measure is reasonable by discussing potential concerns. The main concern could be a simultaneity issue, such as the opposite direction of the cause-and-effect relationships from our theoretical reasoning. This issue would occur when a single firm invests in IT based on having highly specialized workers in the first place. In such cases, our IT progress measure could not be considered to be exogenous. However, as we use Swiss data for our skills and labor market outcomes and German data for our industry-level IT progress, our data should not suffer from such simultaneity

[23] Changes in wages may also occur because individuals e.g., switch firms, but we consider such changes as part of the outcome and therefore do not include them in our regressions because they can be considered as "bad controls" (Angrist and Pischke 2009). In any case, results would not change if we would include them or other such controls in our regressions.

issues.[24] Furthermore, our IT progress variable measures an annual industry-wide growth rate, and it is very unlikely that a single firm's IT investment decision in Switzerland influences the measured IT progress in Germany. Therefore, we argue that our IT progress is likely to be exogenous and does not capture the effect of workers with high levels of specialization or social skills driving our IT progress measure (e.g., because their skills are scarce and therefore drive IT progress). However, given how we measure IT progress, we argue that the potential argument of the skill composition of Swiss workers driving the IT progress in Germany is very unlikely. Therefore, we argue that our assumption that we capture the effect of exogenously driven IT progress is very reasonable. Nevertheless, we acknowledge that having a quasi-experimental design (e.g., a policy change facilitating IT investments with random timing) would provide an even stronger case for exogeneity and causality. Future research could try to use such research design to further investigate our theoretical considerations and empirical findings.

Furthermore, our estimation approach has four additional limitations that occur due to data restrictions. A first limitation is that our skills data from curricula are at the occupational level (in the last paragraph of Sect. 2 on theoretical considerations and hypotheses, we explain how occupations correspond to our theoretical reasoning), and one could argue that observing skills at the *individual* level would be better suited. However, worker-level skill data are not available at such a fine-grained skill level in any administrative dataset. Therefore, we argue that approximating workers' individual skills with a detailed skill extraction from occupational training curricula is even an advantage and an important contribution to closing such individual data gaps.

A second limitation of our empirical approach is that we have to assume that occupational training curricula capture the skills of individuals currently working in a particular occupation (even if they would have been trained in another occupation) as we infer workers' skills from their current occupation. In other words, we assume that if workers are working in an occupation, they have acquired the necessary skills to do so (otherwise, firms would not employ them for a job in this occupation), and these required skills are documented in the occupational training curricula.

A third limitation of our approach arises because we are not able to use firm-level data, so we do not actually observe production processes and IT use in single firms. Thus we cannot directly observe the productivity effects of the interaction between IT use and workers' skills. Instead, we have to assume that individual wage returns sufficiently proxy workers' productivity in firms, which is, however, a fairly common assumption in business or labor economics research. A fourth limitation of our empirical approach, one could argue, is that our IT progress variable only captures the growth of IT capital from computer hardware in Germany because there is no such data for Switzerland. However, we argue that our approach has the advantage of an exogenous IT progress measure and

[24] We take the IT progress measure from Germany, because Germany and Switzerland have similar industry structures and 80% of Switzerland is German speaking. Given that we lack statistical power, we cannot provide language-specific analyses for the other language parts of Switzerland with French or with Italian data. But we conducted a subsample analyses in which we only use the German-speaking part of Switzerland with the German EUklems data and find that our main results are all robust.

that, therefore, this measure is not driven by the skills of the Swiss workers whom we study in our analysis. To summarize, we acknowledge that our empirical approach is limited by the availability of data, but, at the same time, our approach also has major advantages and thus we provide first important empirical insights.

6 Results

Our main results show how IT progress affects the wages of workers in general (see Table 3, column 1) and of workers in occupations with high specialization (hereafter, "specialized workers") and in occupations with high social skills (hereafter, "workers with high social skills") (cf., Table 3, column 5 and the corresponding marginal effects in Fig. 1). Results show that IT progress—on average—leads to positive wage returns if we do not distinguish between workers with different levels of specialization or social skills (cf., Table 3, column 1). However, wage returns differ across workers with different levels of specialization and social skills (cf., column 5 in Table 3, and Fig. 1).

Figure 1 provides insights into the returns for workers with the four different combinations of high or low specialization and high or low social skills. The figure is based on the full interaction model (column 5 in Table 3), and displays the wage returns for an increase in IT progress by 100 percentage points. A main result from the figure is that a ten-percentage-point increase in IT progress leads to a 1.59% increase in wages for *specialized* workers who also have a *high* level of social skills and to a 1.05% increase for *specialized* workers who have only a *low* level of social skills. This means that for a realistic IT growth of 4 percentage points, which roughly corresponds to the mean of years with positive IT growth in the manufacturing sector, wages increased by 0.64% for specialized workers with high social skills and by 0.44% for specialized workers with low social skills. These effects are also economically relevant when we compare the size to e.g., the average annual wage increase in Switzerland, which was, on average, roughly 1.02% (minimum 0.3%, maximum 2.5%) in the relevant time period.[25] Thus IT progress has not only a significant but also a relevant positive effect on wages for workers in specialized occupations—regardless of whether their occupations have high or low levels of social skills. The point estimations of the marginal effects are highest for workers with a combination of high specialization and high social skills, which indicates that high social skills combined with high specialization could be particularly valuable during times of IT progress; however, the difference is not statistically significant (overlapping confidence intervals) thus we refrain from overinterpreting this difference.

In addition, Fig. 1 shows that IT progress does not lead to significant increases in wages for workers with low specialization—regardless of whether their occupations have high or low levels of social skills. None of the marginal effects is statistically

[25] The number of 1.02% is based on a "back-of-the-envelope" calculation from the table "Schweizerischer Lohnindex nach Sektor", taking the geometric mean of the nominal annual changes from 1999–2016 (table accessed on January 8, 2023, https://www.bfs.admin.ch/bfs/de/home/statistiken/arbeit-erwerb/loehne-erwerbseinkommen-arbeitskosten/lohnindex.assetdetail.22304327.html).

Table 3 Wage effects of IT progress for workers below age 45 with high specialization and/or high social skills

Variables	(1) log wage	(2) log wage	(3) log wage	(4) log wage	(5) log wage
IT Progress	0.094* (0.055)	0.004 (0.049)	0.065 (0.043)	– 0.008 (0.048)	0.008 (0.050)
High Specialization*IT Progress		0.136*** (0.033)		0.132*** (0.037)	0.098* (0.053)
High Social Skills*IT Progress			0.048 (0.043)	0.025 (0.036)	– 0.009 (0.053)
High Specialization*High Social Skills*IT Progress					0.063 (0.072)
Controls: age, age^2, year, individual FE	Yes	Yes	Yes	Yes	Yes
Constant	10.633*** (0.403)	10.652*** (0.412)	10.639*** (0.409)	10.655*** (0.413)	10.656*** (0.415)
Observations	55,989	55,989	55,989	55,989	55,989
R^2	0.003	0.004	0.003	0.004	0.004
Number of individuals	29,145	29,145	29,145	29,145	29,145
R^2 overall	0.000732	0.000656	2.02e–05	0.000949	0.000777
R^2 between individuals	0.000995	0.000642	8.32e–05	0.000912	0.000663

Authors' calculations, data based on their own skills measures, EUklems data and the SESAM, 1999–2016, for the sample of workers below age 45. High specialization is constructed after standardizing the specialization measure and taking the values equal and above zero. The same procedure is applied for the construction of high social skills. Standard errors in parentheses clustered on the occupation

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

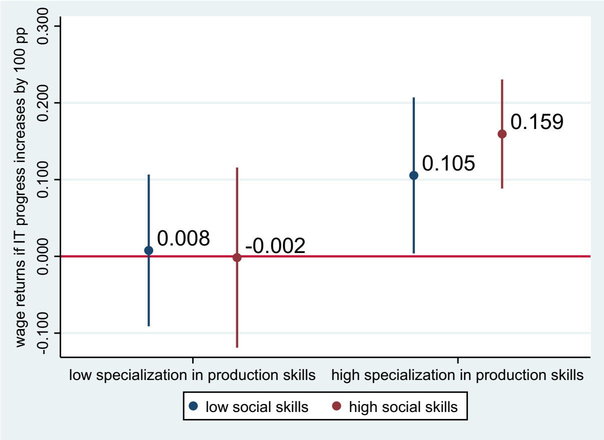


Fig. 1 Marginal effects of regression (5) in Table 3. Authors' calculations, data based on their skills measures, EUklems data and the SESAM, 1999–2016 for the sample of workers below age 45. Marginal effects of regression (5) in Table 3. The same variable construction as in regression (5) apply for this figure. (High specialization is constructed after standardizing the specialization measure and taking the values equal and above zero. The same procedure is applied for the construction of high social skills.)

different from zero. Moreover, taking all the results from Fig. 1 together, we conclude that high levels of social skills are, per se, not valuable during times of IT progress; they only become valuable in combination with a highly specialized occupational skill set. In contrast, the results of Fig. 1 indicate that highly specialized occupational skill sets always benefit during times of IT progress, particularly if they are combined with high levels of social skills.[26,27]

Our main results can be summarized as follows: First, IT progress has a positive effect on wages for workers with high specialization (regardless of whether they have high social skills or not). Second, for high social skills, we find a different pattern than the one for high specialization: IT progress does not have a positive effect on wages for workers with high social skills—unless they are also specialized. The second result is less precisely estimated, but the pattern is still visible. These results are in line with scenario two of our theoretical considerations, i.e., we find support for our hypotheses that IT progress leads to increasing wage returns for workers with high specialization (Hypothesis 1) but not for high social skills (Hypothesis 2b).

[26] In our empirical analyses, we do not differentiate between different workers who are still in their training occupation and workers who changed their occupation because we assume that the latter ones also acquired the skills that are necessary to meet the demands of the occupation they work in. Nevertheless, we conduct additional analyses with a subsample of workers whose current occupation does not match their training occupation. These additional analyses show two results, which are consistent with our main analysis. First, for workers with high specialization (and low social skills), the point estimate on the effect of IT progress is still positive; however, it lacks statistical power due to the small sample size. Second, for workers with *both* high specialization and high social skills, the point estimate on the effect of IT progress is even significantly positive.

[27] Additional analyses using periods with only positive IT growth lead to structurally similar but non-significant results due to the lower number of cases that reduced power.

While our main analyses focus on workers aged below 45, we also report results for workers aged above 44 in Appendix 2 (see Table 6 and Fig. 4),[28] in which we conduct the same estimations as for the younger workers. Results show that IT progress does not have an effect on their wages, also not for specialized vs. nonspecialized workers and not for workers with high vs. low social skills (see Table 6 and Fig. 4).[29] Thus the wages of older workers seem to be generally rather unaffected by IT progress, which is why we in the first place concentrate on wage effects for younger workers and which is consistent with previous research (e.g., Bertschek and Meyer 2009).

7 Conclusion

This paper studies how IT progress and the increasing use of computers in the workplace affect workers with different types of occupational skill sets. Specifically, we examine whether workers with highly specialized skill sets benefit or suffer from IT progress, how workers with high vs. low levels of social skills are affected, and whether a combination of specialized skills and high social skills may be advantageous with increasing IT. To derive empirically testable hypotheses, we provide a theoretical analysis that combines insights from two strands of economic models (Becker and Murphy 1992; Deming 2017) and adds IT progress as an additional input factor. Our theoretical analysis uses Becker and Murphy's (1992) insight that higher specialization increases the productivity of individual workers but also the coordination costs between workers. From Deming (2017), we use the insight that coordination costs may be decreased by social skills because these skills improve communication and collaboration. We add IT progress as another factor to the theoretical analysis, arguing that it potentially also decreases coordination costs because advances in IT facilitate communication and collaboration.

Our theoretical considerations of the interplay of these three factors provide empirically testable hypotheses on the effects of IT progress for workers with different skill sets. Increasing IT equipment at the workplace (e.g., increased availability of computers and email) can be expected to make the coordination between highly specialized workers easier and therefore lead to higher wages for workers with high levels of specialization in comparison to workers with low levels of specialization (H1). For the effect of IT progress on social skills, the theoretical analysis provides two competing hypotheses. On the one side, IT progress could make social skills more valuable because social skills improve coordination between highly specialized workers (H2a). On the other side, increasing IT equipment could make social skills less valuable because communication via IT equipment could replace the need

[28] Previous studies have investigated IT progress and different effects on workers depending on their age (e.g., Aubert et al. 2006 and Bertschek and Meyer 2009), but differences in age are not the focus of our paper.

[29] Our theoretical considerations do not account for differences between age groups. Of course, age groups may differ with respect to occupations, tasks and their exposure to IT progress. Future research may include these differences in their theoretical models.

for social skills to communicate and collaborate (H2b). Thus it is an empirical question, which of the two effects dominates the overall outcome.

To empirically test the hypotheses, we combine three different datasets. In the first dataset, we use natural language processing methods to construct novel proxies to measure the skill sets of workers by using text data from training curricula. We apply this new method to Swiss VET curricula, thereby providing novel data for the levels of specialization and social skills for middle-skilled workers in Switzerland based on their occupation. Second, from the EUklems database, we take data for Germany to proxy IT progress across industries and time (we use German and not Swiss data to particularly avoid endogeneity issues). Third, we use wages and workers' characteristics from administrative labor market data (the Swiss SESAM data).

Our empirical results show that IT progress significantly increases wages for highly specialized workers but not for nonspecialized workers (i.e., support for H1). These results are independent of these workers having high or low social skills. More precisely, a ten-percentage-point increase in IT equipment leads to a 1.59% increase in wages for *highly specialized* workers with *high* social skills and to a 1.05% increase for *highly specialized* workers with *low* social skills. These results are also economically meaningful: For an IT growth of 4 percentage points, which roughly corresponds to the mean of all years with positive IT growth in the manufacturing sector, wages increase by 0.64% for specialized workers with high social skills and by still 0.44% for specialized workers with low social skills. Given an average annual wage increase of 1.02% in Switzerland in the relevant time period, these effects are substantial and, therefore, economically very relevant.

Looking at the empirical results from the perspective of social skills, they show that workers with high levels of social skills do not receive higher wages per se when IT advances (i.e., support for H2b); only if they simultaneously possess a highly specialized occupational skill set, workers with high levels of social skills also receive higher wages. While the latter result is less precisely estimated, our results still indicate that in times of rapid IT progress, a combination of highly specialized occupations and high levels of social skill is particularly useful.

With our results, we add two main contributions to the previous literature. First, we add novel empirical evidence on the returns to more or less specialized occupational skill sets in times of increasing digitalization. Second, we provide evidence on how the effects of specialization and social skills interact in times of digitalization. While research already shows that the returns to social skills increased in the long-run (Deming 2017; Deming and Kahn 2017; Langer and Wiederhold 2023; Weinberger 2014), there has not yet been evidence of how these returns were affected by differing degrees of IT progress. This paper thus adds novel results on how different types of skill sets are affected by the increasing use of IT equipment, i.e., in times of an ever-increasing digital transformation.

The findings also lead to important policy implications. They are relevant for various decision-makers who are responsible for curricula design, such as firms that provide continuous training measures, educational policymakers who decide on occupational curricula, or training institutions (such as technical colleges or universities) that aim at adjusting or reinventing their study programs. While in the context of digitalization, typical policy advice of the past has been to make skill sets more general to increase workers' adaptability and general employability, the empirical findings of our paper indicate that neither specialized skill sets may necessarily be a disadvantage nor general skill sets may automatically be an advantage. On the contrary, our results show that increased availability of IT can make specialized skill sets more valuable because IT progress decreases coordination costs. In addition, our findings indicate that a combination of highly specialized skill sets with high levels of social skills further advances the positive effects of IT progress. Although the latter finding is less precisely estimated and not statistically significant, it is nevertheless important to investigate further in future research.

Last but not least, the paper also makes an important methodological contribution. We introduce a novel method to proxy workers' skill sets in great detail by using novel machine learning methods (based on Eggenberger and Backes-Gellner 2023, and Kiener et al. 2023). Specifically, we show how the application of NLP to curriculum texts can be used to construct new types of skills measures e.g., for occupational specialization or for social skills. Future research can apply these methods to derive proxies for the level of specialization and social skills of basically any other educational curriculum, for example, to college or university curricula. Furthermore, the method can also be used to derive proxies for other types of skills that researchers might be interested in, such as particular types of technical skills, digital skills, or even different types of soft skills. Thus our methodological contribution of measuring novel skills proxies has broad applicability that goes far beyond the research question of this paper.

Appendix 1: mathematical modeling of the relationship between specialization, social skills, IT progress and productivity

This appendix provides a brief version of the mathematical model that backs the theoretical considerations that are presented in a qualitative manner and by illustrative examples in the main part of the paper. The full model is provided in a working paper version in Kiener et al. (2022a) from April 2022.

Our model is based on Becker and Murphy (1992) and Deming (2017) with the additional factor of IT progress. We describe the model step by step, starting with specialization and coordination costs, introducing social skills, and then introducing IT progress.

Specialization The economy produces one good, Y. A continuum of production skills, s, along a unit interval needs to be applied to produce the good. The assumption that all production skills need to be applied to produce the good leads to a Leontief production function as follows (as in Becker and Murphy 1992, who argue that a complementary assumption would lead to similar results):

$$Y = \min_{0 \leq s \leq 1} Y(s) \tag{3}$$

Y: production of one good, s: one particular production skill

The model captures specialization (meaning the more focused the skill set, the more productive the worker is when applying a particular skill) in the output function of the individual worker per skill. This output function is the time invested in production skill s times the productivity of the worker in the production skill s. This productivity increases with time invested in the particular production skills (thereby incorporating specialization).

$$Y_{i,s} = T_{i,s} * P_{i,s}(T_{i,s}) \tag{4}$$

$Y_{i,s}$: output of a worker i per skill s, $T_{i,s}$: worker i's time invested in production skill s, $P_{i,s}$: productivity function of worker i per skill s.

In parallel, given that all workers are identical, the output function per skill overall workers is the same as in formula (4) without the indicator for the individual worker.

$$Y_s = T_s * P_s\left(T_s\right) \tag{5}$$

Y_s: output per skill s.

In team production, each worker's output positively depends on the number of team members that is the same as the number of production skills each worker performs, i.e., the specialization of each worker (thus, the derivate of each worker's output with respect to specialization is positive). Without any costs, the average output per worker looks as follows:

$$Y_i = Y_i(n),\ \text{where} \frac{\partial}{\partial n} Y_i(n) > 0 \tag{6}$$

Y_i: output per worker i, n: specialization.

Due to the complementarity assumption of the Leontieff production function, the worker's output function can also be displayed as follows:

$$Y_i(n) = \frac{1}{n} \min\left(Y_s\right) \tag{7}$$

Coordination costs We now introduce coordination costs that arise in team production. Coordination costs are higher, the more workers in a team, i.e., the more specialized the workers (thus, the derivate of coordination costs with respect to specialization is positive).

$$C_i(n), \text{where} \frac{\partial}{\partial n} \mathrm{C_i(n)} > 0 \tag{8}$$

C_i: coordination costs per worker i, n: specialization.

Social Skills To facilitate teamwork, workers can devote time investing in social skills ($\mathrm{T_z}$). Each team member has one-unit time that they can devote to either production skills or social skills (in line with Becker and Murphy 1992, but social skills added).

$$1 = T_{i,s} + T_{i,z} \tag{9}$$

$T_{i,s}$: worker i 's time invested in production skills, $T_{i,z}$: worker i 's time invested in social skills.

To continue with the model, we introduce now an assumption that is line with Becker and Murphy (1992): An efficient team concentrates on an equal *set of production skills*, $w = 1/n$.

Taken the assumption and formula (9) together, the output per production skill s is:

$$Y_s = T_s * w = T_s * \frac{1}{n} = 1 - T_z \tag{10}$$

w: set of production skills, T_s: time invested in production skills, T_z: time invested in social skills

Solving formula (10) for the time spent on a particular production skill leads to:

$$T_s = n - n * T_Z \tag{11}$$

Substituting (19) into formula (5) $Y_s = T_s * P_s(T_s)$ results in a more detailed formula for output per production skill. The part of this output function that shows productivity per skill increases with specialization and decreases with social skills:

$$Y_s = Y_s\left(n, T_Z\right) = \left(n - n * T_Z\right) * P_s\left(n, T_Z\right),\ where\ \frac{\partial}{\partial n}P_s\left(n, T_Z\right) > 0, \frac{\partial}{\partial T_Z}P_s\left(n, T_Z\right) < 0 \tag{12}$$

Given formula (7), the output per worker also depends on social skills, thus $Y_i\left(n, T_Z\right)$, and following from (12), the output per worker decreases in social skills (thus the first derivative of the output with respect to social skills is negative):

$$Y_i\left(n, T_Z\right), \text{where} \frac{\partial}{\partial n}\mathrm{Y_i}\left(\mathrm{n}, T_Z\right) > 0, \frac{\partial}{\partial T_Z}\mathrm{Y_i}\left(\mathrm{n}, T_Z\right) < 0 \tag{13}$$

Finally, social skills work similar to a depreciation factor to coordination costs. The depreciation factor is always between 0 and 1. The depreciation factor decreases when social skills increase (thus the derivate of the depreciation factor with respect to social skills is negative). Taken together, social skills decrease the coordination costs.

$$Y_i\left(n, T_Z\right), \text{where} \frac{\partial}{\partial n}\mathrm{Y_i}\left(\mathrm{n}, T_Z\right) > 0, \frac{\partial}{\partial T_Z}\mathrm{Y_i}\left(\mathrm{n}, T_Z\right) < 0 \tag{14}$$

Welfare optimization The formulas up to now build the foundation for the welfare function that depends on specialization and social skills. The welfare per worker consists of output minus depreciated costs.

$$W_i = Y_i\left(n, T_Z\right) - \delta\left(T_Z\right) * C_i(n) = (n - n * T_Z) * P_s(n, T_Z) - \delta\left(T_Z\right) * C_i(n) \tag{15}$$

Welfare contains the two parameters, specialization and social skills. Thus welfare is optimized along these two parameters leading to two optimization formulas. First, the first derivative of the welfare function with respect to specialization equals 0. Second, the first derivative of the welfare function with respect to social skills equals 0.

$$\frac{\partial}{\partial n}W_i = \frac{\partial}{\partial n}Y_i(n, T_Z) - \delta(T_Z) * \frac{\partial}{\partial n}C_i(n) = 0,\ where\ \frac{\partial}{\partial n}C_i(n) > 0 \tag{16}$$

$$\frac{\partial}{\partial T_Z}W_i = \frac{\partial}{\partial T_Z}\mathrm{Y_i}(\mathrm{n}, T_Z) - \mathrm{C_i(n)}\frac{\partial}{\partial T_Z}\delta(T_Z) = 0,\ where\ \frac{\partial}{\partial T_Z}\delta(T_Z) < 0,\ \frac{\partial}{\partial T_Z}\mathrm{Y_i}(\mathrm{n}, T_Z) < 0 \tag{17}$$

The two optimization formulas in (16) and (17) can be displayed in one formula. To do so, we need three intermediary steps:

From formula (17), we know:

$$\mathrm{C_i(n)} = \frac{1}{\frac{\partial}{\partial T_Z}\delta(T_Z)}\frac{\partial}{\partial T_Z}\mathrm{Y_i}(\mathrm{n}, T_Z) \tag{18}$$

The deviation of formula (18) with respect to specialization is:

$$\frac{\partial}{\partial n}\mathrm{C_i(n)} = \frac{1}{\frac{\partial}{\partial T_Z}\delta(T_Z)}\frac{\partial}{\partial T_Z \partial n}\mathrm{Y_i}(\mathrm{n}, T_Z) \tag{19}$$

Then we substitute formula (19) into the rearranged formula (16):

$$\frac{\partial}{\partial n}\mathrm{Y_i}(\mathrm{n}, T_Z) = \delta(T_Z) * \frac{1}{\frac{\partial}{\partial T_Z}\delta(T_Z)} * \frac{\partial}{\partial T_Z \partial n}\mathrm{Y_i}(\mathrm{n}, T_Z) \tag{20}$$

The final formula of welfare optimization looks as follows:

$$\frac{\frac{\partial}{\partial n}\mathrm{Y_i}(\mathrm{n}, T_Z)}{\frac{\partial}{\partial T_Z \partial n}\mathrm{Y_i}(\mathrm{n}, T_Z)} = \frac{\delta(T_Z)}{\frac{\partial}{\partial T_Z}\delta(T_Z)} \tag{21}$$

IT progress and welfare optimization IT affects the welfare function through the depreciation factor of coordination costs. The depreciation factor decreases when IT increases (thus the first derivative of the depreciation factor with respect to IT is negative). The new function of the depreciation factor looks as follows:

$$\delta = \delta(T_Z, IT),\ \frac{\partial}{\partial T_Z}\delta(T_Z, IT) < 0, \frac{\partial}{\partial IT}\delta(T_Z, IT) < 0 \tag{22}$$

IT: Information technologies (i.e., a continuous input that is higher, the more IT).

Therefore, the welfare function including IT looks as follows:

$$W_i = Y_i(n, T_Z) - \delta(T_Z, IT) * C_i(n) = n * (1 - T_Z) * P_s(n, T_Z) - \delta(T_Z, IT) * C_i(n) \tag{23}$$

The final welfare optimization formula is the same as previously (formula 21) but now the depreciation factor also depends on IT:

$$\frac{\frac{\partial}{\partial n}\mathrm{Y}_{\mathrm{i}}\left(\mathrm{n},T_Z\right)}{\frac{\partial}{\partial T_Z \partial n}\mathrm{Y}_{\mathrm{i}}\left(\mathrm{n},T_Z\right)}=\frac{\delta\left(T_Z,IT\right)}{\frac{\partial}{\partial T_Z}\delta\left(T_Z,IT\right)} \tag{24}$$

Scenarios The model decribed above provides two opposing scenarios of how IT progress affects the optimal level of specialization and social skills. While in both scenarios, IT progress reduces coordination costs, leading to an increase of the optimal level of specialization, the two scenarios differ in their optimal level of social skills: In scenario one, the optimal level of social skills increases, because IT progress leads to a higher optimal level of specialization, which in turn increases the need for coordination, thereby increasing the demand for social skills (similar to an "income effect"). In scenario two, the "income effect" described in scenario one is cancelled out by a "substitution effect": Because IT progress makes coordination more efficient, one unit of social skills will now lead to a higher coordination cost reduction. Thus, for a given level of specialization, less social skills are required in the optimum, leading to decreasing demand for social skills. This decrease is a consequence of the trade-off between investing in social skills and investing in production skills. Investing less time in social skills increases the time available to invest in production skills, thereby increasing the output. If social skills are used more efficiently through IT, the optimal level of social skills decreases (i.e., IT substitutes social skills).

To emphasize that the model indeed provides the two scenarios, we show in the subsequent subsection that—given a particular, reasonable functional form of the welfare function—IT progress can lead to an increase or a decrease in the optimal level of social skills. In contrast, the optimal level of specialization increases in both scenarios. In our example, we show that both scenarios can occur. In scenario one, the "income effect" dominates, and in scenario two, the "substitution effect" dominates.[30]

Specific examples of functions in the model This subsection shows specific examples of the welfare function to emphasize that the model indeed shows these scenarios. In our examples, the welfare function (output minus depreciated costs) incorporates two assumptions, one on the depreciation factor and one on the output function. The assumptions are as follows: The depreciation factor contains IT in a nonlinear, such as exponential, way interacting with social skills (T_z). In our example, the depreciation factor is $\frac{1}{(IT*T_z)^4}$. The output function contains a strong punishment for time invested in social skills; in other words, output decreases strongly with less time invested in production skills (i.e., specialization n). Thus, in our example, the output function includes a log function for this strong punishment and looks as

[30] In the following subsection, we show how the functional form of the welfare function would look, such that the optimal level of social skills can increase or decrease. In particular, we incorporate specific assumptions into the welfare function (output minus depreciated costs). The assumptions concern both the depreciation factor and the output function. We assume that the depreciation factor includes IT in a nonlinear (e.g., exponential) manner through interacting with social skills. We also assume that the output function contains a strong punishment for time invested in social skills (i.e., output strongly decreases with less time invested in production skills). These assumptions lead to a welfare function in which both scenarios are possible, depending on the level of the IT increase.

follows $(n - \log(n * T_z + 0.1)) * (n - n * T_z)$. In sum, our welfare function (output minus depreciated costs) contains a simple coordination cost function ($\frac{n^3}{100}$), and the two assumptions on the depreciation factor and the output function:

$$\begin{aligned} W &= output - depreciation\,factor * coordination\,costs \\ &= \left(n - \log\left(n * T_z + 0.1\right)\right) * \left(n - n * T_z\right) - \frac{1}{\left(IT * T_z\right)^4} * \frac{n^3}{100} \end{aligned} \tag{25}$$

To show the two scenarios, we start with the baseline that has no IT progress, then continue with the first scenario with a small increase in IT (optimal level of specialization and the one of social skills increases compared to the baseline), and finally show the second scenario with a high increase in IT (optimal level of specialization increases and the optimal level of social skills decreases compared to the baseline). We take the graphs from wolframalpha.com, x corresponds to n (i.e., specialization) in the model, and t corresponds to T_z (i.e., social skills) in the model (see Table 4).

Having shown with some specific examples that the scenarios are indeed shown by our theoretical model, we proceed by presenting the empirically testable hypotheses.

Empirically Testable Hypotheses As described in Sect. 2 in the main paper and given the theoretical model described in this appendix, we can derive empirically testable hypotheses. H1 applies to both scenarios, H2a applies to the first scenario and H2b applies to the second scenario. The hypotheses look as follows:

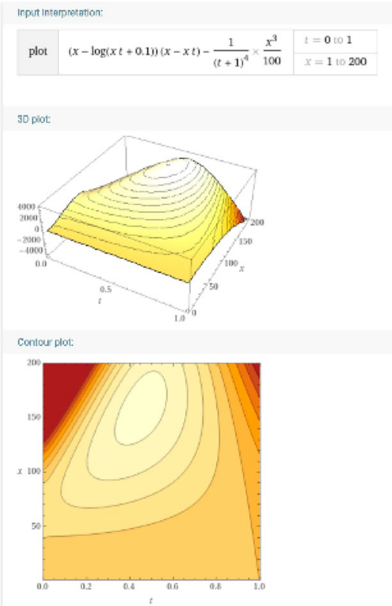
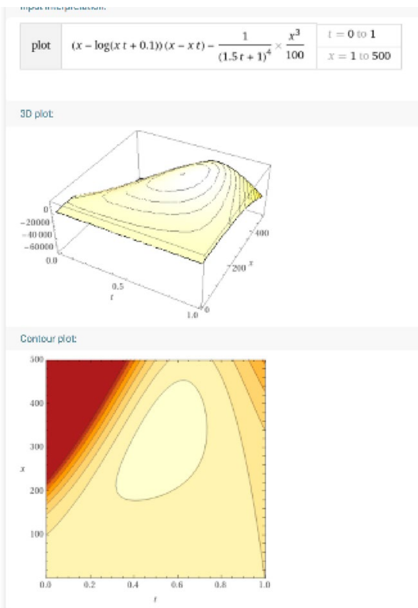
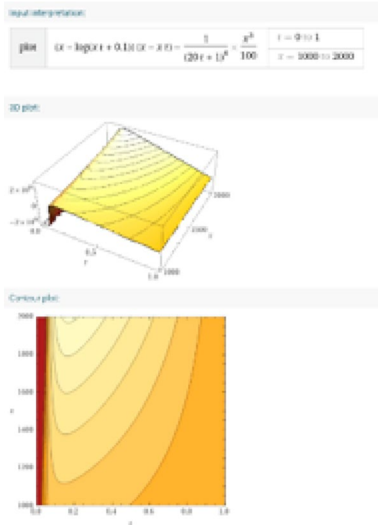
(H1) Increasing IT equipment in the workplace makes the coordination between specialized workers easier and therefore makes more specialization beneficial to the worker

(H2a) Increasing IT equipment makes the social skills of workers more valuable (because with increasing IT a higher specialization of workers is more efficient and—as a complement—workers also require more social skills to better communicate = scenario 1).

(H2b) In contrast, increasing IT equipment could also substitute social skills—instead of complementing them—and in this case would not increase the value of social skills (because communication via IT replaces the need for workers' social skills = scenario 2).

These empirically testable hypotheses are the basis for our empicial analyses in the main paper.

Table 4 The example of our welfare function in the baseline, scenario 1, and scenario 2

Baseline	Scenario 1: Small <u>IT progress</u>	Scenario 2: High <u>IT progress</u>
plot (x-log(x*t+0.1))(x-x*t)-(1/(t+1)^4)*(x^3/100), t=0…1,x=1…200	plot (x-log(x*t+0.1))(x-x*t)-(1/(**<u>1.5</u>**t+1)^4)*(x^3/100), t=0…1,x=1…500	plot (x-log(x*t+0.1))(x-x*t)-(1/(**<u>20</u>**t+1)^4)*(x^3/100), t=0…1,x=1000…2000
		
t* around 0.45 x* around 170	t* around 0.55 → increases compared to baseline x* around 300 → increases compared to baseline	t* goes to 0 but stops before → decreases compared to baseline x* indef. → increases compared to baseline

Authors' calculations with graphs from wolframalpha.com

Appendix 2: Additional figures and tables

See Figs. 2, 3, 4 and Table 5, 6, 7.

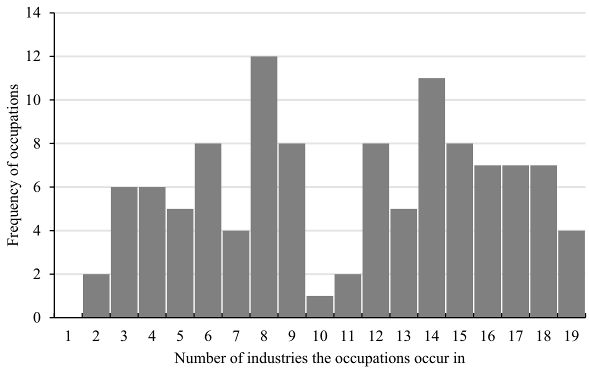


Fig. 2 Frequencies of occupations across different industries. Authors' calculations of how many times each of the 111 occupations occur in different industries (19 in total), data based on their skills measures, the EUklems data and the SESAM, 1999–2016

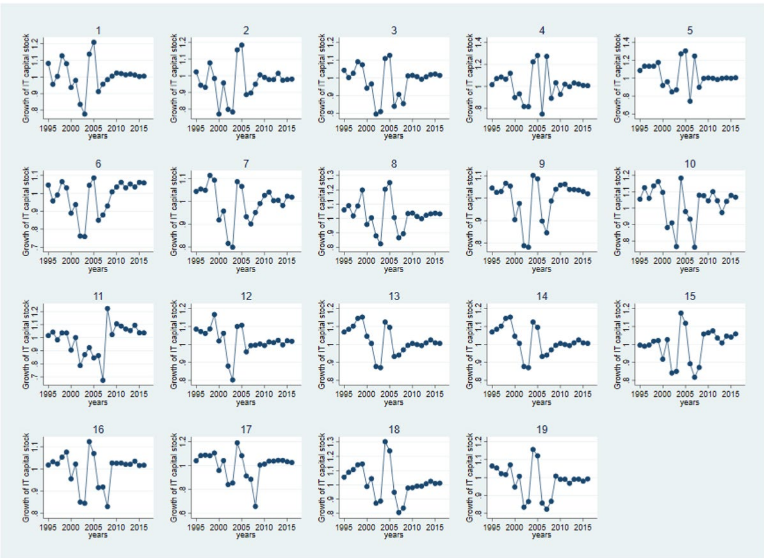


Fig. 3 Variation of growth of IT capital stock across industries. Authors' calculations of our IT progress variable (Growth of IT capital stock) from the EUklems data, 1995–2017. The industry codes are as follows: 1: (A) Agriculture, forestry and fishing, 2: (B) Mining and quarrying, 3. (C) Manufacturing, 4. (D) Electricity, gas, steam and air conditioning supply, 5. I Water supply; sewerage; waste management and remediation activities, 6. (F) Construction, 7. (G) Wholesale and retail trade; repair of motor vehicles and motorcycles, 8. (H) Transportation and storage, 9. (I) Accommodation and food service activities, 10. (J) Information and communication, 11. (K) Financial and insurance activities, 12. (L) Real estate activities, 13. (M) Professional, scientific, and technical activities, 14. (N) Administrative and support service activities,15. (O) Public administration and defense; compulsory social security, 16. (P) Education, 17. (Q) Health and social work, 18. I Arts, entertainment and recreation, 19. (S) Other service activities

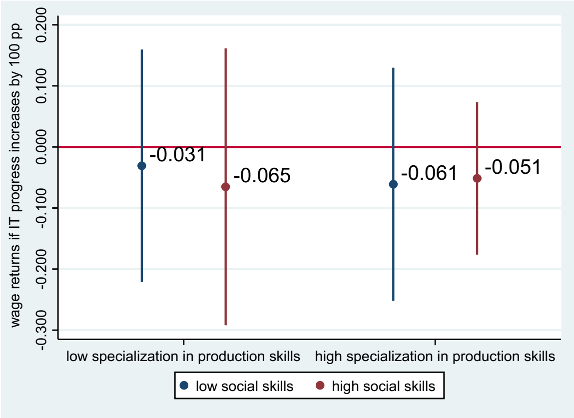


Fig. 4 Marginal effects of regression (5) in Table 7. Authors' calculations data, based on their skills measures, EUklems data and the SESAM, 1999–2016, for the sample of workers above age 44. Marginal effects of regression (5) in Table 7. The same variable construction as in regression (5) apply for this figure. (High specialization is constructed after standardizing the specialization measure and taking the values equal and above zero. The same procedure is applied for the construction of high social skills.)

Table 5 Correlations between wages, specialization, and social skills

	Wages	IT Progress	Specialization (continuous)	Social Skills (continuous)
Wages				
IT Progress	− 0.0065			
Specialization (continuous)	−0.0072	0.0148*		
Social Skills (continuous)	−0.0658*	−0.0095	−0.0737*	

Authors' calculations data based on their skills measures, the EUklems data and the SESAM, 1999–2016, for the sample of workers below age 45. High specialization is constructed after standardizing the specialization measure and then taking the values equal and above zero. Same procedure is applied for the construction of high social skills. Pairwise correlations with a Bonferroni correction over all years. * means statistical significance on the 0.01 significance level

Table 6 Summary statistics of wages and IT progress by high/low specialization and high/low social skills

	Variable	Mean	St. dev	Min	Max	N
low specialization + low social skills	Annual wages	58,984	28,642.25	50	574,500	15,089
	IT progress	0.98	0.096	0.66	1.30	15,089
high specialization + low social skills	Annual wages	54,233	26,561.29	50	354,700	11,694
	IT progress	0.99	0.10	0.66	1.30	11,694
low specialization + high social skills	Annual wages	53,849	26,179.8	100	321,900	10,780
	IT progress	0.98	0.10	0.66	1.30	10,780
high specialization + high social skills	Annual wages	52,555	31,489.56	100	984,700	18,426
	IT progress	0.98	0.11	0.66	1.30	18,426

Authors' calculations of the summary statistics over all individuals by groups high/low specialization and high/low social skills from the sample, data based on their skills measures, EUklems data and the SESAM, 1999–2016, for the sample of workers below age 45. High specialization is constructed after standardizing the specialization measure and taking the values equal and above zero. The same procedure is applied for the construction of high social skills. Annual wages are not calculated in full-time equivalents because we want to include lower wages arising, for example, due to part-time work

Table 7 Wage effects of IT progress for workers above age 44 with high specialization and/or high social skills

Variables	(1)	(2)	(3)	(4)	(5)
	log wage	log wage	log wage	log wage	log wage
IT Progress	– 0.052 (0.073)	– 0.046 (0.090)	– 0.046 (0.088)	– 0.042 (0.096)	– 0.031 (0.097)
High Specialization*IT Progress		– 0.009 (0.060)		– 0.008 (0.063)	– 0.030 (0.084)
High Social Skills*IT Progress			– 0.011 (0.057)	– 0.010 (0.060)	– 0.034 (0.109)
High Specialization*High Social Skills*IT Progress					0.044 (0.124)
Controls: age, age^2, year, individual FE	Yes	Yes	Yes	Yes	Yes
Constant	0.256 (1.108)	0.256 (1.108)	0.255 (1.108)	0.254 (1.108)	0.253 (1.108)
Observations	41,165	41,165	41,165	41,165	41,165
R^2	0.023	0.023	0.023	0.023	0.023
Number of individuals	21,467	21,467	21,467	21,467	21,467
R^2 overall	2.69e–05	3.75e–05	2.99e–05	3.82e–05	8.32e–05
R^2 between individuals	7.60e–07	3.26e–09	6.16e–07	2.21e–09	1.20e–05

Authors' calculations, based on their skills measures, EUklems data and the SESAM, 1999–2016, for the sample of workers above age 44. High specialization is constructed after standardizing the specialization measure and taking the values equal and above zero. The same procedure is applied for the construction of high social skills. Standard errors in parentheses clustered on the occupation

Acknowledgements This paper has benefitted from many audiences and exchanges. In particular, we would like to thank the anonymous referees for the helpful comments as well as the following individuals for their help and support: Eric Bettinger, Thomas Dohmen, Simon Janssen, Uli Kaiser, Harald Pfeifer, Jens Mohrenweiser, Dinand Webbink, conference participants at SASE, at LISER-IAB conference, at SFUVET VET Congress 2022, at PILLARS, at COPE 2022, at First International Leading House Conference on the "Economics of Vocational Education and Training", at the Annual Meeting of the German Economic Association (Verein für Socialpolitik), as well as seminar participants at the University of Zurich for helpful comments and and Natalie Reid for language consulting. We thank the Swiss Federal Statistical Office for providing the labor market data.

Author contributions FK: conceptualization, data curation, methodology and formal analysis, writing (draft & editing); C.E.: data curation, funding acquisition, methodology and formal analysis, writing (draft, review, & editing); UB-G: conceptualization, funding acquisition, supervision, writing (draft, review, & editing). All authors have contributed fair shares, with FK having the largest contribution and CE the second largest. The order of authors listed in the manuscript has been approved by all authors.

Funding Open access funding provided by University of Zurich. This study is funded partly by the SNSF within the framework of the national research program "Digital Transformation" (NRP 77) and partly by the Swiss State Secretariat for Education, Research, and Innovation (SERI) through its "Leading House VPET-ECON: A Research Center on the Economics of Education, Firm Behavior and Training Policies".

Availability of data and materials The SESAM dataset was made available from the Swiss Federal Statistical Office under a data usage contract. EUklems data is freely and openly available online at https://euklems.eu/archive-history/download-archive/ (Stehrer et al. 2019). The skills datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

Angrist J, Pischke J-S (2009) Mostly harmless econometrics. An empiricist's companion. Princeton Univ. Press, Princeton

Aubert P, Caroli E, Roger M (2006) New technologies, organisation and age: firm-level evidence. Econ J 116 (509), F73-F93. http://www.jstor.org/stable/3590484. Accessed 26 May 2021

Becker GS, Murphy KM (1992) The division of labor, coordination costs, and knowledge. Q J Econ 107(4):1137–1160. https://doi.org/10.2307/2118383

Bertschek I, Meyer J (2009) Do older workers lower IT-enabled productivity? Jahrbücher Für Nationalökonomie Und Statistik. https://doi.org/10.1515/jbnst-2009-2-315

Borghans L, ter Weel B (2006) The division of labour, worker organisation, and technological change. Econ J 116(509):F45–F72. https://doi.org/10.1111/j.1468-0297.2006.01064.x

Cameron CA, Miller DL (2015) A practitioner's guide to cluster-robust inference. J Human Resours 50(2):317–372. https://doi.org/10.3368/jhr.50.2.317

Campello RJGB, Moulavi D, Sander J (2013) Density-based clustering based on hierarchical density estimates. In: Pei J, Tseng VS, Cao L, Motoda H, Xu G (eds) Advances in knowledge discovery and data mining. In: 17th Pacific-Asia Conference, PAKDD 2013, Gold Coast, Australia, April 14–17, 2013, Proceedings, Part II, vol. 7819. 1st ed. 2013. Berlin, Heidelberg: Springer Berlin Heidelberg; Imprint: Springer (Lecture Notes in Artificial Intelligence, 7819), pp 160–172

Carbonell J, Goldstein Jade (1998) The use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries. In : Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval: Association for Computing Machinery (SIGIR '98), pp 335–336

Caroli E, van Reenen J (2001) Skill-biased organizational change? Evidence from a panel of British and French Establishments. Q J Econ 116(4):1449–1492. https://doi.org/10.1162/003355301753265624

Deming D (2017) The growing importance of social skills in the labor market. Q J Econs 132(4):1593–1640. https://doi.org/10.1093/qje/qjx022

Deming D, Kahn LB (2017) Skill requirements across firms and labor markets: evidence from job postings for professionals. J Labor Econs 36(S1):S337–S369. https://doi.org/10.1086/694106

Dessein W, Santos T (2006) Adaptive organizations. J Polit Econ 114(5):956–995. https://doi.org/10.1086/508031

Devlin J, Chang M-W, Lee K, Toutanova K (2018) BERT: pre-training of deep bidirectional transformers for language understanding. https://arxiv.org/pdf/1810.04805.

Eggenberger C, Backes-Gellner U (2023) IT skills, occupation specificity and job separations. Econ Educ Rev 92:102333. https://doi.org/10.1016/j.econedurev.2022.102333

Eggenberger C, Rinawi M, Backes-Gellner U (2018) Occupational specificity: a new measurement based on training curricula and its effect on labor market outcomes. Labour Econ 51:97–107. https://doi.org/10.1016/j.labeco.2017.11.010

Eggenberger C, Janssen S, Backes-Gellner U (2022) The value of specific skills under shock: High risks and high returns. Labour Econ 78:102187. https://doi.org/10.1016/j.labeco.2022.102187

Gaggl P, Wright GC (2017) A short-run view of what computers do: evidence from a UK tax incentive. Am Econ J Appl Econ 9(3):262–294. https://doi.org/10.1257/app.20150411

Gordon R, Sayed H (2019) The industry anatomy of the transatlantic productivity growth slowdown. Cambridge, MA (National Bureau of Economic Research Working Paper Series, 25703).

Green A (2002) The many faces of lifelong learning: recent education policy trends in Europe. J Educ Policy 17(6):611–626. https://doi.org/10.1080/0268093022000032274

Jansen A, de Grip A, Kiener B (2017) The effect of choice options in training curricula on the demand for and supply of apprentices. Econ Educ Rev 57:52–65. https://doi.org/10.1016/j.econedurev.2017.02.003

Kiener F, Eggenberger C, Backes-Gellner Uschi (2022a) How IT Progress affects Returns to Specialization and Social Skills (Swiss Leading House "Economics of Education" Working Paper, 192).

Kiener F, Gnehm A-S, Clematide S, Backes-Gellner U (2022b) IT skills in vocational training curricula and labour market outcomes. J Educ Work 35(6–7):614–640. https://doi.org/10.1080/13639080.2022.2126968

Kiener F, Gnehm A-S, Backes-Gellner U (2023) Noncognitive skills in training curricula and nonlinear wage returns. Int J Manpower. https://doi.org/10.1108/IJM-03-2022-0119

Langer C, Wiederhold S (2023) The value of early-career skills (CESifo Working Paper Series, 10288). Available online at https://www.cesifo.org/en/publications/2023/working-paper/value-early-career-skills. Accessed 25 Feb 2023

Lindbeck A, Snower DJ (2000) Multitask learning and the reorganization of work: from tayloristic to holistic organization. J Labor Econ 18(3):353–376. https://doi.org/10.1086/209962

Marsden D (1999) A theory of employment systems. In: Micro-foundations of societal diversity. Oxford University Press, Oxford

Reimers N, Gurevych I (2019) Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. https://arxiv.org/pdf/1908.10084. Accessed 19 Jun 2021

Rinawi M, Backes-Gellner U (2021) Labour market transitions after layoffs: the role of occupational skills. Oxf Econ Papers 73(1):76–97. https://doi.org/10.1093/oep/gpz064

Stehrer R, Bykova A, Jäger K, Reiter O, Schwarhappel M (2019) Industry level growth and productivity data with special focus on intangible assets (wiiw Statistical Report No. 8). https://euklems.eu/archive-history/download-archive/. Accessed 15 Feb 2021

Weinberger C (2014) The increasing complementarity between cognitive and social skills. Rev Econ Stat 96(5):849–861

Yi F, Jiang Bo, Wu J (2020) Topic modeling for short texts via word embedding and document correlation. IEEE Access 8:30692–30705. https://doi.org/10.1109/ACCESS.2020.2973207

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Fabienne Kiener[1] · **Christian Eggenberger**[1] · **Uschi Backes-Gellner**[1]

✉ Uschi Backes-Gellner
backes-gellner@business.uzh.ch

Fabienne Kiener
fabienne.kiener@business.uzh.ch

[1] Department of Business Administration, University of Zurich, Plattenstrasse 14, 8032 Zurich, Switzerland