INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 860

Learning Entrepreneurship as an Employee

Yoko ASUYAMA*

September, 2022

Abstract

Most entrepreneurs learn entrepreneurial skills while working as employees. Thus, employees' opportunities to learn entrepreneurial skills are expected to substantially influence the economy's entrepreneurship rate and entrepreneurial performance. However, previous studies have assessed such learning opportunities using indirect, rough measurements. This study utilizes a more direct and comprehensive measure. I first analyze entrepreneurs' tasks by using information on 47 tasks from 31 countries and find that entrepreneurs perform more autonomous and diverse tasks, financial and managerial tasks, and fewer clerical tasks than employees. Next, using individual-level data from 23 countries in 2012–2017, I demonstrate that employees' greater learning opportunities for entrepreneurial tasks increase individuals' self-perceived entrepreneurial skills and their probability of becoming an entrepreneur. These relationships are generally robust to alternative learning opportunity indices, instrumental variable estimations, omitted variable bias, and learning-environment-level regressions. My bound estimates imply that a 1-standard-deviation increase in an employee's opportunities to learn entrepreneurial tasks increases the probability of an individual becoming an entrepreneur by 0.5–1.5 percentage points, which is equivalent to 4%–12% of the mean entrepreneurship rate. By contrast, having more opportunities to learn entrepreneurial tasks does not result in a higher level of employment or innovativeness in entrepreneurs' businesses.

Keywords: Entrepreneur, Entrepreneurial skills, Entrepreneurship, GEM, PIAAC

JEL classification: J24, L26, M13, M50

---

* Associate Senior Research Fellow, Business and Industry Studies Group, Development Studies Center, IDE (Yoko_Asuyama@ide.go.jp)

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

Learning Entrepreneurship as an Employee*

Yoko Asuyama†

Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO)

September 2022

Abstract

Most entrepreneurs learn entrepreneurial skills while working as employees. Thus, employees' opportunities to learn entrepreneurial skills are expected to substantially influence the economy's entrepreneurship rate and entrepreneurial performance. However, previous studies have assessed such learning opportunities using indirect, rough measurements. This study utilizes a more direct and comprehensive measure. I first analyze entrepreneurs' tasks by using information on 47 tasks from 31 countries and find that entrepreneurs perform more autonomous and diverse tasks, financial and managerial tasks, and fewer clerical tasks than employees. Next, using individual-level data from 23 countries in 2012–2017, I demonstrate that employees' greater learning opportunities for entrepreneurial tasks increase individuals' self-perceived entrepreneurial skills and their probability of becoming an entrepreneur. These relationships are generally robust to alternative learning opportunity indices, instrumental variable estimations, omitted variable bias, and learning-environment-level regressions. My bound estimates imply that a 1-standard-deviation increase in an employee's opportunities to learn entrepreneurial tasks increases the probability of an individual becoming an entrepreneur by 0.5–1.5 percentage points, which is equivalent to 4%–12% of the mean entrepreneurship rate. By contrast, having more opportunities to learn entrepreneurial tasks does not result in a higher level of employment or innovativeness in entrepreneurs' businesses.

Keywords: Entrepreneur, Entrepreneurial skills, Entrepreneurship, GEM, PIAAC

JEL codes: J24, L26, M13, M50.

---

* I thank all the internal seminar participants at the Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO). I am grateful for the funding support from the JSPS KAKENHI [Grant Number JP19J00295] and the IDE-JETRO. More than half of the research were conducted when I was the JSPS Research Fellow at the Waseda University.

† E-mail: Yoko_Asuyama@ide.go.jp

1. Introduction

Entrepreneurs and entrepreneurship are important sources of innovation, economic growth, and employment growth (Parker 2018). Most entrepreneurs accumulate work experience and obtain ideas for new businesses while working as an employee (Sørensen and Fassiotto 2011; Bhidé 2000). For example, Burton et al. (2002) examine the career history of the founders of 164 young high-technology firms in Silicon Valley in the mid-1990s and find that over 90% of them worked for established employers before launching their businesses. Based on the labor market transition matrices created by Boeri et al. (2020: Table 4)—who utilize national-level labor force surveys from the United States (US), the United Kingdom (UK), and Italy—95% (US), 90% (UK), and 59% (Italy) of self-employed persons who changed their employment status in 2017 were employees in 2016.[1] Provided that employees account for 85%–94% of all workers in advanced countries,[2] the amount of entrepreneurial skills and knowledge that employees acquire is expected to substantially influence the economy's entrepreneurship rate and entrepreneurial performance.

This study focuses on this issue by exploring the effect on entrepreneurship of employees' opportunities to learn entrepreneurial tasks. I first analyze the kinds of tasks that entrepreneurs usually perform based on information about 47 tasks from 31 countries. I find that entrepreneurs perform more autonomous and diverse tasks, financial and managerial tasks, and fewer clerical tasks than employees. Next, using individual-level data from 23 countries, I empirically demonstrate that individuals working in an environment in which employees have more opportunities to learn entrepreneurial tasks, have more self-perceived entrepreneurial skills and are more likely to become entrepreneurs. An

---

[1] The remaining percentages were for the unemployed in 2016. These statistics only capture the transition from 2016 to 2017. It is possible that some of the unemployed had been working as employees before 2016.

[2] The data were calculated from the Organisation for Economic Co-operation and Development (OECD) statistics (https://stats.oecd.org/Index.aspx?DataSetCode=ALFS_SUMTAB, accessed March 25, 2022). The ratio of employees (i.e., non-self-employed workers) to total employment in 2020 was 94% in US, 85% in the European Union (27 countries), and 90% in Japan.

individual's learning environment is defined by the same county × gender × age group × education group as that person. In this environment, employees are considered to have more opportunities to learn entrepreneurial tasks when they perform tasks that are similar to those of entrepreneurs. I also examine the influence of employees' learning opportunities on the employment level and the innovativeness of entrepreneurial businesses and find no robust effects.

Many studies have determined that work experiences increase a person's likelihood of becoming an entrepreneur as well as their entrepreneurial performance (Unger et al. 2011; Parker 2018: 157-162). These studies usually assess work experience according to general experience, industry-specific experience, and functional experience (e.g., managerial, marketing, finance, and production experience) (Unger et al. 2011). However, these measurements are often not comprehensive and are either too crude or too specific. Empirical studies based on Lazear's (2005) "Jack-of-all-trades" hypothesis usually examine more comprehensive work experiences, but their focus is on the effect of experience diversity (Wagner 2006; Spanjer and van Witteloostuijn 2017). In contrast, this study identifies the entrepreneurs' task set from data covering a wide range of tasks and countries. It then determines the work experience necessary for an entrepreneur's business according to how much employees perform a similar task set. This measurement is more objective and comprehensive than those used in the existing research.

This study is most closely related to several studies that have examined the entrepreneurship effect of one's environment for learning entrepreneurial skills. These studies measure this environment according to the presence of current or former entrepreneurs in a person's colleagues or family (Nanda and Sørensen 2010; Lindquist et al. 2015), the entrepreneur ratio in the same social group (Giannetti and Simonov 2009), firm density in the resident region (Guiso et al. 2021), the number of layers in a firm (Tåg et al. 2016), a firm's worker diversity (Marino et al. 2012), and a country's degree of aging (Liang et al. 2018). These measures are indirect and rough in capturing the environment for learning entrepreneurial skills. It remains unknown what skills or knowledge individuals learn in these environments. By contrast, the current study measures entrepreneurial tasks directly by utilizing comprehensive information on 47 tasks. As explained in Section 2, my analytical framework is based on Guiso et al. (2021). However, their empirical proxy for learning opportunities—regional firm density—captures not only learning from entrepreneurs but also non-pecuniary entrepreneurial benefits (e.g., social norms making it desirable to be an entrepreneur) and entrepreneurs' informal credit market networks, as discussed by Ginnetti and Simonov (2009). By contrast, it is unlikely that my learning opportunity measurement captures these alternative non-learning channels.

The remainder of this paper is organized as follows. Section 2 presents the analytical and empirical framework and the main data sources. Section 3 defines an entrepreneur and analyzes entrepreneurial tasks. Section 4 provides the estimation results for the effect of employees' opportunities to learn entrepreneurial tasks on skill acquisition and occupational choice. This includes several robustness checks and discusses implications using Japan as an example. Section 5 provides the estimation results for the effects on entrepreneurial performance. Section 6 concludes.

2. Analytical Framework and Main Data Sources

2.1 Learning Opportunities and Entrepreneurship: Theory

Guiso et al. (2021) present a learning opportunity model, which was originally presented by Guiso and Schivardi (2011), who modified Lucas's (1978) occupational choice model. In their model, opportunities to learn entrepreneurial skills that vary by environment ("location" in their model) affect individual occupational choices through two channels. In one channel, having more learning opportunities improves an individual's entrepreneurial skills (the skill improvement effect). In the other channel, having more learning opportunities reduces the entry cost of starting an entrepreneurial business (the entry cost reduction effect). Both effects unambiguously increase the probability of a person becoming an entrepreneur.

By contrast, the total effect of having more learning opportunities on the average entrepreneur's performance is theoretically ambiguous and an empirical question. This is because although the skill improvement effect generally works positively, the entry cost reduction effect enables less skilled individuals to become entrepreneurs and thus decreases the average entrepreneur's performance.

2.2 Empirical Approach and the Main Data Sources

In this subsection, I briefly explain my empirical approach and main data sources. A more detailed explanation of the empirical method is provided in Sections 3–5. I first analyze entrepreneurial tasks (in comparison to employee tasks) using information on 47 tasks from 31 countries by applying exploratory factor analysis (Section 3.2). I use data from the Public Use Files of the Programme for the International Assessment of Adult Competencies (PIAAC) conducted by the Organisation for Economic Co-operation and Development (OECD).[3] PIAAC provides cross-national, harmonized information about key cognitive skills and various workplace tasks of adults aged 16–65 (OECD 2016). I use data for 2011–2012 (first cycle, round 1), which includes 157,567 adults from 23 countries, and data for 2014–2015 (first cycle, round 2), which includes 43,021 adults from eight countries.

Second, I construct an employee-level learning opportunity index ($EntrLearn_i$) that measures how similar the tasks of employee i are to those of entrepreneurs (Section 3.3). This index is computed as the predicted probability that employee i is an entrepreneur, judging only from i's tasks, after controlling for other individual characteristics. Third, an employee's average $EntrLearn_i$ is computed for each learning environment, which is defined by county $\times$ gender $\times$ age group $\times$ education group (*cgae*) cell (Section 4.1). $EntrLearn_{cgae}$ is the main variable in my study, which measures employees' opportunities to learn entrepreneurial tasks in the *cgae* environment.

Fourth, I merge $EntrLearn_{cgae}$ with individual-level entrepreneurial activity data using *cgae* as an identifier. The entrepreneurial activity data are taken from the Adult

[3] The data are downloaded from the OECD's PIAAC website (https://www.oecd.org/skills/piaac/data/, accessed September 2019). The German PIAAC Scientific Use File (Rammstedt et al. 2016) is also used to obtain more detailed data for Germany.

Population Survey dataset of the Global Entrepreneurship Monitor (GEM).[4] Since 1998, GEM has provided cross-national, harmonized datasets of entrepreneurship activity that have been used widely in various studies (Reynolds et al. 2005; Álvarez et al 2014; GEM 2022). I use entrepreneurship activity and individual characteristic data for 2012–2017, which include 570,742 individuals from 30 countries. By utilizing individual-level data from 23 countries that match $EntrLearn_{cgae}$ (constructed only from the countries surveyed in 2011–2012), I regress self-perceived entrepreneurial skills or an entrepreneur indicator on $EntrLearn_{cgae}$. This tests whether individuals working in an environment in which employees have more opportunities to learn entrepreneurial tasks improve their entrepreneurial skills and increases their probability of becoming an entrepreneur as theory predicts (Section 4.2). I also conduct robustness checks using alternative learning opportunity indices, performing instrumental variable (IV) estimation, computing bound estimates robust to omitted variable bias, and performing learning-environment-level regressions (Sections 4.3–4.6). Finally, I estimate the effects of employees' learning opportunities on entrepreneurs' performance; that is, employment level and innovativeness of entrepreneurial businesses (Section 5).

3. Entrepreneurial Tasks and Employees' Learning Opportunities

3.1 Definition of an Entrepreneur in PIAAC Data

In this section, I utilize both rounds' PIAAC data, including 31 countries, to ensure that the sample has a sufficient number of observations for examining the tasks of entrepreneurs.[5] I first define an entrepreneur as a self-employed worker who has

---

[4] The data are downloaded from the Global Entrepreneurship Monitor (GEM) website (https://www.gemconsortium.org/data/sets?id=aps, accessed June 2019–June 2021). Regarding the codebook, see the GEM website and Reynolds (2021).

[5] The first-round PIAAC countries include Austria, Belgium (Flanders only), Canada, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, Norway, Poland, Russia (excluding the Moscow municipal area), the Slovak Republic, South Korea, Spain, Sweden, the UK (England and Northern Ireland only), and the US. The second-round countries include Chile, Greece, Israel, Lithuania, New Zealand, Singapore, Slovenia, and Turkey.

employees working for him or her. In other words, I exclude solo self-employed workers.[6] This is because solo self-employed workers are likely to capture underemployment rather than entrepreneurship (Boeri et al. 2020). Recent studies define an entrepreneur as an incorporated self-employed person (Levine and Rubinstein 2017) or an individual starting a limited liability company (Tåg et al. 2016). PIAAC data do not distinguish whether a self-employed person is incorporated or not. However, by examining the National Longitudinal Survey of Youth in the US, Levine and Rubinstein (2017: Table I) report that the median number of employees is 0.0 for the unincorporated self-employed, whereas it is 2.0 for the incorporated self-employed. Thus, my definition of an entrepreneur is likely to overlap with the incorporated self-employed. Furthermore, PIAAC scores regarding tasks that often involve co-workers in the same organization are either missing or very low in the case of solo self-employed workers. Thus, to ensure the comprehensiveness of the tasks, the solo self-employed must be excluded from entrepreneurs.

When using PIAAC data, I do not restrict "entrepreneurs" to early-stage ones (who have launched their enterprise within the past 3.5–5 years), whereas I do so when examining entrepreneurial activity using GEM data. This is because it is necessary (i) to restrict entrepreneurial activities to those that occurred around or after the PIAAC surveys and (ii) to ensure a sufficient number of entrepreneur observations in PIAAC.

3.2 Entrepreneurial Tasks

I assume that entrepreneurial skills can be learned on the job by performing tasks similar to those of entrepreneurs. What entrepreneurial skills are necessary to become an entrepreneur? The OECD and the European Union describe entrepreneurship skills as "*a combination of technical skills, business management skills, and personal skills required for starting and operating in business and self-employment. For example, they include*

[6] According to the PIAAC background questionnaire, a self-employed worker includes those "who have their own business or are partners in a business as well as freelancers." In addition, "[a] self-employed person may or may not have personnel," which includes regular employees as well as "family members working paid or unpaid in the business."

team building, negotiation, strategy development, financial planning, and marketing" (OECD and EU 2019: 318). The Harvard Business School's free e-book[7] identifies the following seven skills as necessary to become an innovative entrepreneur: (i) basic financial skills such as budgeting and financial statement analysis; (ii) networking; (iii) the ability to accept and act on feedback; (iv) pattern recognition when examining data such as financial statements, sales, and market data; (v) strategic thinking, including analytical, communication, problem-solving, planning, and management skills; (vi) negotiation; and (vii) a growth mindset. Hartog et al. (2010) find that mathematical, social, and technical abilities are more valuable for entrepreneurs, whereas verbal and clerical abilities are more valuable for employees. Other studies claim that not only specific skills but also a certain skill mix is essential. Lazear (2004, 2005) and several subsequent empirical studies (e.g., Wagner 2006; Aldén et al. 2017) show that entrepreneurs require balanced skillsets; that is, an entrepreneur must be a so-called "Jack of all trades."

I identify entrepreneurial tasks from comprehensive task information data. I utilize PIAAC's 47 questions (measured on a 5-point scale) on various workplace tasks related to management, problem-solving, learning, information and communication technology (ICT) use, reading, writing, numerical work, and clerical work (Table A1). It includes most tasks related to the aforementioned entrepreneurial skills. Regarding the 42 questions that ask for the frequency of the tasks, I assign a more accurate average frequency (per week): I assign 0 for the original frequency "1 = *never*," 0.12 for 2 = *less than once a month*, 0.62 for 3 = *less than once a week but at least once a month*, 3 for 4 = *at least once a week but not every day*, and 5 for 5 = *every day*. Regarding the one question on the amount of time spent cooperating or collaborating with co-workers, I assign the midpoint time ratio: I assign 0 for 1 = *none of the time*, 0.125 for 2 = *up to a quarter of the time*, 0.375 for 3 = *up to half of the time*, 0.750 for 4 = *more than half of the time*, and 1 for 5 = *all the time*. The remaining four questions assess the extent of the respondent's work autonomy (1 = *not at all*, 2 = *very*

[7] *Harvard Business School Online*, "So You Want to Be an Entrepreneur: How to Get Started" (https://info.online.hbs.edu/entrepreneurship-ebook?_ga=2.113852039.1895745335.1651020907-580554233.1651020907, accessed on February 3, 2022).

little, 3 = *to some extent*, 4 = *to a high extent*, and 5 = *to a very high extent*). To analyze these 47 task measures with a mixture of different scales, I next standardize each of them based on the entire sample of workers. Table A2 compares the average of these 47 standardized task scores for entrepreneurs and employees.[8] It shows that the scores of most autonomy- and finance-related tasks and some of the ICT- and management-related tasks (e.g., those related to e-commerce, planning others' activities, and negotiations) are particularly higher for entrepreneurs than for employees.

In the next step, I perform an exploratory factor analysis (EFA) on these 47 standardized task scores based on 6,676 entrepreneur observations. The purpose of EFA is to reduce the number of variables (i.e., task items) and find a few common meaningful factors that contain most of the original information. This data reduction process is essential for analyzing the overall picture of entrepreneurial tasks and to avoid multicollinearity between the tasks in the regression analysis. I apply the principal factor method with varimax orthogonal rotation, which is one of the most common techniques (Costello and Osborne 2005). Orthogonal rotation, which generates factors that are uncorrelated with each other, is chosen primarily to avoid multicollinearity between factors in the regression analysis. The factors obtained by oblique rotation, which allows correlation between factors, are used in the robustness check.

I extract five factors based on the following criteria: (i) a scree plot indicates four to six factors; (ii) the traditional Kaiser's rule, which recommends retaining factors with an eigenvalue greater than 1, indicates five factors; and (iii) each extracted factor should have a meaningful interpretation. The five factors are named as follows (with their abbreviations in parentheses): ICT; typical managerial work and learning (*Mgmt*); autonomy; finance and sales (*Finance*); and clerical and analytical work (*Clerical*). Table A3 reports the factor loadings of the 47 task items. I do not drop any tasks with low factor loadings (i.e., loadings less than 0.3 or 0.4) or those with loadings relatively high for two factors. This

[8] According to the PIAAC background questionnaire, an employee is defined as "someone who gets a salary or wage from an employer or a temporary employment agency."

goes against the usual EFA conventions. However, low-loading or multi-loading tasks do not mean that they are not important for an entrepreneur's work. In fact, tasks related to problem-solving and negotiation, which are identified as important entrepreneurial skills in the literature, are dropped if I drop low-loading and multi-loading items. Thus, in the main analysis, I calculate the factor scores by keeping all 47 task items to maintain the comprehensiveness of workplace tasks.[9] The task factor scores generated after dropping low-loading and multi-loading task items are used in the robustness check.

Based on Lazear's Jack-of-all-trades hypothesis that a balanced skillset is essential for entrepreneurs, I also construct a task diversity index (*TaskDiversity*) as follows[10]:

$$TaskDiversity_i = 1 - \sum_{j=1}^{J} (x_{ij}/\sum_{j=1}^{J} x_{ij})^2, \quad (1)$$

where subscripts i and j represent worker and task, respectively. x_{ij} is i's unstandardized approximated frequency of task j. Only 42 task items assessing task frequency are used. Thus, J is 42. This index measures task diversity because it is calculated as 1 minus the Herfindahl–Hirschman Index, which is a measure of concentration.

I also construct the following alternative task diversity measure (*TaskDiversityF*):

$$TaskDiversityF_i = 1 - \sum_{g=1}^{G} (x_{ig}/\sum_{g=1}^{G} x_{ig})^2. \quad (2)$$

The 42 task frequency items are first classified into factor-based group g according to the highest-loading factor in Table A3. G is five, the number of factor-based groups. x_{ig} is the average of the unstandardized approximated frequency of all tasks classified into group g.[11] *TaskDiversityF* prevents the task diversity measure from being driven by tasks in a particular area. In the main analysis, I use *TaskDiversity*, not *TaskDiversityF*, because the former is less correlated with the five task factor scores and thus creates less multicollinearity in the subsequent regression analysis. *TaskDiversityF* is used in the robustness check.

---

[9] To calculate the factor scores, the regression scoring method, which provides the highest correlation between the factor scores and the estimated factors (DiStefano et al. 2009), is applied.

[10] Asuyama (2022) uses the same task diversity index as a job enlargement measure.

[11] To construct x_{ig}, the frequency of *manuwork*, which negatively loads on the ICT factor, is reversed.

Table 1 compares the values of the five task factor scores, two task diversity indices, and some basic characteristics for entrepreneurs and employees. For comparison, the task factor scores are standardized with a mean of 5 and a standard deviation (SD) of 1.[12] All statistics are based on the regression sample in the next subsection. This shows that on average, entrepreneurs perform more autonomous work, finance- and sales-related tasks, and managerial and learning tasks, than employees do. Although the difference is much smaller, entrepreneurs also perform more ICT tasks. By contrast, entrepreneurs perform fewer clerical and analytical tasks than employees. Both *TaskDiversity* and *TaskDiversityF* demonstrate that entrepreneurs perform more diverse tasks than employees.

3.3 Measuring Employees' Opportunities to Learn Entrepreneurial Tasks

As the next step, I construct an index ($EntrLearn_i$) that measures how similar employee i's tasks are to those of entrepreneurs; that is, how much employee i is performing (and thus learning) entrepreneurial tasks. I compute $EntrLearn_i$ as the predicted probability that employee i is an entrepreneur, judging only from i's tasks, after controlling for other individual characteristics.

I first estimate the following probit model using all PIAAC observations for entrepreneurs and employees:

$$\mathrm{P}(Y = 1|\mathbf{Z}) = \Phi(\alpha_0 + \mathbf{TaskF}\boldsymbol{\beta_0} + \gamma_0 TaskDiversity + \mathbf{X}\boldsymbol{\delta_0}), \qquad (3)$$

where Φ is the standard normal cumulative distribution function. Y is 1 if the respondent is an entrepreneur and 0 if the respondent is an employee. **Z** denotes the full set of explanatory variables. **TaskF** is a vector of five task factor scores (i.e., *ICT*, *Mgmt*, *Autonomy*, *Finance*, and *Clerical*). *TaskDiversity* is the task diversity index, as described in Section 3.2. **X** is a vector of respondent characteristics: this includes a female dummy, age and its square, various ability-related measures (i.e., years of education, literacy and numeric proficiency scores, years of work experience and its square, mother's education,

[12] The mean is set to 5, so that all task factor scores do not have a negative value. A negative value is inconvenient in the subsequent regression analysis, particularly when adding the squared terms of the task scores.

father's education, number of books at home, three factor scores for tasks performed outside of work), learning attitude, health status, household size, living with a partner dummy, having children dummy, foreign-born dummy, foreign native language dummy, indices for trust and altruism, various work-related measures (i.e., working hours, dummies for public sector and non-profit organization sector, occupation dummies [nine categories], industry dummies [16 categories]), and country dummies. Detailed explanations of these variables are provided in Table A4. All regressions apply the senate weights that reflect PIAAC's sampling weights within a country but give each country equal total weight.

Then, $EntrLearn_i$ is computed as the predicted probability that employee *i* is an entrepreneur, judging only from *i*'s tasks ($\mathbf{TaskF}_i$ and $TaskDiversity_i$):

$$EntrLearn_i = \Phi(\widehat{\alpha_0} + \mathbf{TaskF}_i\widehat{\boldsymbol{\beta}_0} + \widehat{\gamma}_0 TaskDiversity_i + \overline{\mathbf{X}}\widehat{\boldsymbol{\delta}_0}), \qquad (4)$$

where *i* is an employee, and $\widehat{\alpha_0}$, $\widehat{\boldsymbol{\beta}_0}$, $\widehat{\gamma}_0$, and $\widehat{\boldsymbol{\delta}_0}$ are estimated from Equation (3). $\overline{\mathbf{X}}$ is the average of $\mathbf{X}$ in the regression sample.

Table 2 reports the probit regression results for Equation (3). In the main analysis, I use the estimated coefficients and $\overline{\mathbf{X}}$ in Column (1) to compute $EntrLearn_i$. Consistent with Table 1, employees are judged to be more like entrepreneurs if they perform more *Autonomy*, *Finance*, followed by *Mgmt* and *ICT* tasks, fewer *Clerical* tasks, and more diversified tasks. Similar trends are observed with fewer control variables using a higher number of observations (Column 2), with the alternative task diversity index (*TaskDiversityF*, Column 3), or with the squared terms of the task variables added (Column 4), although the average marginal effects (AME) of *ICT* become negative in the last case.

Figure 1 reports the average five factor scores and *TaskDiversity* for entrepreneurs and employees who are separated by the $EntrLearn_i$ quintile. It confirms that $EntrLearn_i$ approximates task similarity with entrepreneurs because scores are generally becoming closer to those of entrepreneurs as $EntrLearn_i$ becomes higher (Panel a). Similar graphs are obtained for $EntrLearnSQ_i$, which is the probability that employee *i* is an entrepreneur predicted from *TaskDiversity*, five task factor scores, as well as their squared terms (Panel b).

4. Results: Employees' Learning Opportunities and Entrepreneurship

4.1 Baseline Empirical Approach

To examine the relationship between employees' opportunities to learn entrepreneurial tasks and entrepreneurial activity, I primarily consider the following model:

$$Y^*_{icgaet} = \alpha_1 + \beta_1 EntrLearn_{cgae} + \mathbf{X}_{icgaet}\boldsymbol{\gamma}_1 + \varepsilon_{icgaet}, \quad (5)$$

where subscripts *i*, *c*, *g*, *a*, *e*, and *t* denote working person (*i*), country (*c*), gender (*g*: male or female), age group (*a*: either age 16–24, 25–34, 35–44, 45–54, and 55–65), education group (*e*: either up to post-secondary non-tertiary education, or tertiary education and higher), and year (*t*: 5 years from 2012 to 2017), respectively. Y^*_{icgaet} is either (i) self-perceived level of entrepreneurial skill, (ii) net utility of being an entrepreneur, or (iii) entrepreneurial performance. In most cases, Y^*_{icgaet} is a latent variable, and only the binary indicator $Y_{icgaet} = 1[Y^*_{icgaet} > 0]$ is observable. Y^*_{icgaet} and Y_{icgaet} are taken from GEM data. $EntrLearn_{cgae}$ is the average value of employees' $EntrLearn_i$ in a cell (i.e., a learning opportunity environment) defined by country × gender × age group × education group (*cgae*) computed from PIAAC data.[13] It is crucial that $EntrLearn_{cgae}$ is constructed from an employee sample that does not include entrepreneurs. $\mathbf{X}_{icgaet}$ is a vector of the control variables that includes *icgaet*-, *cgaet*- and *cgae*-level variables and is constructed from either GEM, PIAAC, or the World Values Survey and European Values Study data (WVS-EVS; World Values Survey Association 2015; European Values Study Foundation 2011), as I discuss in the following section and in Table A5. ε_{icgaet} is the error term.

In the baseline analysis, I estimate either the linear probability model (LPM) or the probit model when Y^*_{icgaet} is unobservable. When Y^*_{icgaet} is observable, I estimate Equation (5) by least squares. All regressions apply the senate weights that reflect the GEM's sampling weights within a country but give each country equal total weight.

I only use the $EntrLearn_{cgae}$ of the PIAAC first-round countries. Consequently,

[13] When there are fewer than 20 observations in a *cgae* cell, $EntrLearn_{cgae}$ is not computed and is therefore excluded from the analysis.

23 countries covering 444 *cgae* environments, which are surveyed by both the GEM (2012–2017) and the first round of PIAAC (2011–2012) are included in the regression analysis. This includes Austria, Belgium, Canada, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, Norway, Poland, Russia, the Slovak Republic, South Korea, Spain, Sweden, the UK, and the US. The summary statistics of the variables are presented in Table 3.

4.2 Skill Acquisition and Occupational Choice: Baseline Results

Because both (i) the self-perceived level of entrepreneurial skill and (ii) the net utility derived from being an entrepreneur are unobservable, I use a binary indicator as the dependent variable. As an indicator for (i), I use a binary response ($Suskill_{icgaet}$) that takes the value of 1 if the respondent answers "yes" to the GEM question "*Do you have the knowledge, skill, and experience required to start a new business?*" Regarding an occupational choice indicator for (ii), I construct two measures from GEM data: the first is $EntrepY5_{icgaet}$, which takes the value of 1 if *i* is either a so-called "nascent entrepreneur" who is involved in setting up a business but has not paid wages for the last 3 months or an "owner-manager of a new firm" that has been operating for less than 5 years. The second measure is TEA_{icgaet}, which takes the value of 1 if *i* is involved in GEM-defined "total early-stage entrepreneurial activity" (*TEA*); that is, either a nascent entrepreneur or an owner-manager of a new firm that has been operating for 3.5 years or less.[14] Owner-managers of a new firm that has been operating for more than 5 years (or those that have been operating for more than 3.5 years) are excluded from the sample of the $EntrepY5_{icgaet}$ (TEA_{icgaet}) regressions. By matching the time-invariant variable $EntrLearn_{cgae}$, which was constructed from PIAAC data for 2011–2012 and entrepreneur indicators that vary across 2012–2017, I assume that employees' opportunities to learn

[14] The years of operation are counted from the year in which wage payments were started. For a more detailed definition of a nascent entrepreneur, an owner-manager of a new firm, and TEA, see the GEM's website (https://www.gemconsortium.org/wiki/1149, accessed on February 8, 2021).

entrepreneurial tasks in a *cgae* environment do not change between 2007 and 2017.[15]

Table 4 reports the baseline LPM and probit regression results. The columns vary by the control variables.[16] Column (1) controls for *icgaet*-level covariates including a female dummy, age and its square, education level (seven categories), household size, a dummy for knowing "*someone personally who started a business in the past 2 years*" (*Knowent*), a dummy for agreeing that "*fear of failure would prevent you from starting a business*" (*Fearfail*), year dummies, and country dummies. Column (2) additionally controls for *cgae*-level variables, which include an average employee's learning attitude, three factor scores for tasks performed outside of work, and *EntrOld*, which is the *cgaet*-level population ratio of old entrepreneurs (owner-managers of a new firm founded more than 10 years before year *t*). *Knowent* and *EntrOld* control for learning entrepreneurial skills from other entrepreneurs (Guiso et al. 2021), non-pecuniary entrepreneurial benefits, or entrepreneurs' informal credit market networks (Giannetti and Simonov 2009). Column (3) further controls for *cgae*-level covariates including the ratio of full-time employees (*Fulltime*), the ratio of employees with an indefinite employment contract (*Permanent*), the average number of employees at the workplace (*Estsize*), and workers' average tenure (*Tenure*), which is a labor mobility indicator. *Fulltime* and *Permanent* are added to capture the possibility that better working conditions prevent employees from leaving an organization and becoming an entrepreneur. *Estsize* captures the "small firm effect" in which employees of small firms are more likely to become entrepreneurs (Parker 2018: 69-70, 222-224). Column (4) further controls for *icgaet*-level variables, including perceived local opportunities for starting a business (*Opport*), perceived attitude of citizens toward the desirability and status of being an entrepreneur

[15] When using $EntrepY5_{icgaet}$ of t = 2012 (or t = 2017), individual i may have become an entrepreneur between 2007 and 2012 (or 2012 and 2017). Regressing $EntrepY5_{icgaet}$ on $EntrLearn_{cgae}$, which is computed using data for 2011–2012 (although data for 2014–2015 are used when estimating the coefficients used to predict $EntrLearn_i$), I implicitly assume that the opportunities to learn entrepreneurial tasks in the *cgae* environment are the same in 2007 and 2017.

[16] Note that I cannot control *i*'s industry affiliation because only entrepreneurs are asked their industry affiliation in the GEM data.

(*Nbgoodc* and *Nbstatus*), and the frequency of media coverage of entrepreneurial success (*Nbmedia*). Column (5) additionally controls for *cgae*-level average worker psychological traits such as trust, locus of control, focus on creativity, need for achievement, and risk preference, most of which are considered to influence entrepreneurship in the literature (Guiso et al. 2006; Parker 2018: 178-196). These trait variables are computed as the average score of the *cgae* cell using the fifth and sixth waves (covering 2004-2016) of the pooled WVS-EVS data. Note that in Columns (3)–(5), several countries are dropped from the analysis because some of their covariates are missing.[17]

Table 4 shows that an employee's opportunity to learn entrepreneurial tasks (*EntrLearn*) is significantly positively associated with the perceived acquisition of entrepreneurial knowledge and skill (*Suskill*) and the occupational choice to become an entrepreneur *(EntrepY5* and *TEA*) in all specifications.[18] Such positive associations are consistent with the theoretical predictions of Guiso et al. (2021) (see Section 2.1). These associations are robust even when considering positive effects from other entrepreneurs by controlling for *Knowent* and *EntrOld*. The size of the coefficients of *EntrLearn* is much greater in Column (5). This is mainly because there is a much smaller sample size in Column (5). If I perform an LPM regression in Columns (1)–(4) based on the same sample as Column (5), the coefficients of *EntrLearn* become much larger (2.792–3.445 for *Suskill* regressions, 2.610–3.303 for *EntrepY5* regressions, and 2.567–3.201 for *TEA* regressions).

4.3 Alternative Measures for Learning Opportunities

Significantly positive associations are also obtained when alternative learning opportunity indices are used (Table A7). These alternative indices include the following: (i) $EntrLearnSQ_{cgae}$, which is predicted from the five task factor scores and *TaskDiversity* as

---

[17] Additional controls dropped Austria, Canada, and US in Column (3), the Czech Republic and Denmark in Column (4), and Belgium, Ireland, Italy, and Slovakia in Column (5). Furthermore, additional controls in Column (4) substantially reduced the number of observations in Belgium, Estonia, Finland, Norway, Poland, Russia, Slovakia, and Sweden.

[18] The results for the control variables (in Columns 2 and 5 in the *Suskill* and *EntrepY5* regressions) are reported in Table A6.

well as the squared terms of these task variables; (ii) $EntrLearnF_{cgae}$, which uses *TaskDiversityF* instead of *TaskDiversity* in Equations (3)-(4); (iii) $EntrLearnOB_{cgae}$, which uses alternative task factor scores extracted using quartimin oblique rotation, which allows correlation between the factors; (iv) $EntrLearnL_{cgae}$, which uses alternative task factor scores that are extracted through the same EFA method as $EntrLearn_{cgae}$ but by dropping low-loading and multi-loading task items; and (v) $EntrLearn40_{cgae}$, which is predicted from alternative task factor scores computed based on 40 task items excluding tasks of a different nature (i.e., four autonomy-related and three learning-related task items) as well as the autonomy index and *TaskDiversity*. Appendix B provides additional detail on the construction of (iii)–(v). The baseline *EntrLearn* and the alternative indices are highly correlated, with correlation coefficients of approximately 0.85–1.00.[19]

4.4 Instrumental Variable (IV) Estimation

The coefficients of $EntrLearn_{cgae}$ are biased when $EntrLearn_{cgae}$ and the error term ε_{icgaet} is correlated in Equation (5). The use of *cgae*-level (instead of individual *i*-level) learning opportunities constructed from a different dataset (PIAAC) mitigates this endogeneity.[20] Furthermore, individuals generally cannot choose aspects of their environment such as country, gender, and age group. Persons with a preference for entrepreneurship may choose a certain education group (e.g., choose to receive a tertiary education to become an entrepreneur). However, controlling for a person's education level breaks the endogeneity resulting from self-selection. Another concern is omitted variable bias: opportunities to learn entrepreneurial tasks that are not observed from the PIAAC dataset may be included in the error term and positively correlated with $EntrLearn_{cgae}$.

[19] Based on *EntrepY5* regression sample in Column (2) of Table 4. The correlation coefficients between the main *EntrLearn* and its alternative indices are 0.970 (*EntrLearnSQ*), 0.995 (*EntrLearnF*), 1.000 *(EntrLearnOB*), 0.854 (*EntrLearnL*), and 0.983 (*EntrLearn40*).

[20] Due to the nature of data, it is not possible to examine the effects of *i*'s *own* learning experiences of entrepreneur tasks. The measurement error resulting from using *cgae*-level learning opportunity generates attenuation bias: the coefficient of $EntrLearn_{cgae}$ becomes biased toward zero. Thus, from the viewpoint of measurement error, the positive coefficient of $EntrLearn_{cgae}$ can be interpreted as the lower-bound estimate for the effects of *i*'s own learning opportunities.

In this case, the coefficient of $EntrLearn_{cgae}$ is likely to be upwardly biased.

To deal with possible endogeneity, I instrument $EntrLearn_{cgae}$ with two *cgae*-level variables. The first is $Boss_{cgae}$, constructed from PIAAC data, which is measured according to the ratio of bosses to all employees in a *cgae* cell. A boss is defined as a person who has one or more subordinates. The second is $BossDistance_{cgae}$, which approximates the non-boss employee's expected "distance" to becoming a boss. It is computed using the following steps. First, using PIAAC employee observations, I calculate the average age of a boss ($BossAvgAge_{cge}$) in a *cge* cell (country × gender × education cell). Second, I compute $BossDistance_{icgae}$, which is calculated as the population counts between the age of non-boss employee *i* and $BossAvgAge_{cge}$ of the same *cge* cell as *i*, divided by the population aged 16–65 in *i*'s resident country. I assume that a greater value for $BossDistance_{icgae}$ indicates that employee *i* has a longer distance to becoming a boss. The data for the population counts by age in each country are taken from the US Census Bureau's International Database (IDB).[21] Lastly, $BossDistance_{cgae}$ is computed as the average of $BossDistance_{icgae}$ in each *cgae* cell.

An entrepreneur is one's own boss and the boss of his or her employees. Thus, I assume that employees perform more entrepreneur-like tasks as they become closer to the position of boss. I assume that both instruments are valid: they affect *i*'s entrepreneurial activity only by affecting employees' tasks, i.e., $EntrLearn_{cgae}$.

In Panel (a) of Table 5, I estimate the LPM by two-stage least squares (2SLS). The results indicate that the instruments are not weak: the first-stage *F* statistic is always greater than 10, the rule-of-thumb cutoff for weak instruments, and is usually greater than 19.93, which is the Stock–Yogo critical value for 10% maximal IV size (Stock and Yogo 2005). The overidentification tests also indicate that my instruments are valid under the assumption that at least one instrument is valid. The coefficient of *EntrLearn* is always positive and greater than that of the baseline LPM in Table 4. It is also always statistically

[21]The data are downloaded from the IDB website (https://www.census.gov/data-tools/demo/idb/, accessed on November 24, 2021).

significant whenever exogeneity of *EntrLearn* is not rejected.

Panel (b) of Table 5 reports the AMEs obtained from the two-stage probit estimation results. I follow the control function approach described in Wooldridge (2010: 586-589): In the first stage, similar to 2SLS, *EntrLearn* is regressed on two instruments and all of the control variables by least squares. In the second stage, I perform a probit regression of *Suskill* (or *EntrepY5* or *TEA*) on *EntrLearn*, the control variables, and the residuals obtained from the first-stage regression, and compute the AME of *EntrLearn*. If the coefficient of the first-stage residual is significantly different from zero, the null hypothesis that *EntrLearn* is exogenous is rejected. To compute the correct standard errors of the AME, I bootstrap the entire process 250 times.[22] The coefficients of the first-stage residuals imply that *EntrLearn* is likely to be endogenous, particularly in the *EntrepY5* and *TEA* regressions. Like in the 2SLS LPM case, the estimated AME is always statistically significant, positive, and greater than that of the baseline probit model in Table 4.

4.5 Bound Estimates Robust to Omitted Variable Bias

Despite the overidentification test results, my instruments may still have some direct effect on entrepreneurial activity and be correlated with unobservables in the error term. Considering the possibility of invalid instruments, I compute the bound estimates for *EntrLearn*'s coefficient (β) in the baseline LPM that are robust to omitted variable bias. I follow Oster's (2019) approach. Oster (2019) shows that the bias-adjusted β, named β^*, can be approximated as follows[23]:

$$\beta^*(R_{max}, \delta) \approx \tilde{\beta} - \delta[\dot{\beta} - \tilde{\beta}] \frac{R_{max} - \tilde{R}}{\tilde{R} - \dot{R}}, \qquad (6)$$

[22] I confirm that the number of replications (250) is sufficient by first experimenting with five different random-number seeds and confirming that the standard errors for the AMEs do not change significantly.

[23] Equation (6) is based on the assumption that the relative contribution of each observed control (*OC*) to *EntrLearn* equals that to *Y*. However, it closely approximates β^* without imposing this special assumption in many cases (Oster 2019: 193). The results in Table 6 do not impose this special assumption and therefore involve more complicated calculation of β^* described in Oster (2019: 193).

where $\tilde{\beta}$ and $\tilde{R}$, respectively, are the coefficients of *EntrLearn* and the R-squared value obtained from a regression of entrepreneurial activity (*Y*) on *EntrLearn* and the observed controls (*OC*). $\dot{\beta}$ and $\dot{R}$, respectively, are the coefficients of *EntrLearn* and the R-squared obtained from a regression of *Y* on *EntrLearn* only. R_{max} is the R-squared obtained from a hypothetical regression of *Y* on *EntrLearn* and both *OC* and unobservable controls (*UC*). δ is $Cov(UC, EntrLearn)/Var(UC)$ divided by $Cov(OC, EntrLearn)/Var(OC)$, where *Cov*(*A, B*) indicates the covariance of A and B, and *Var*(*A*) indicates the variance of A. δ measures "the relative degree of selection on observed and unobserved controls" (Oster 2019: 188); that is, the relative importance of *UC* to *OC* in relation to *EntrLearn*. To compute $\beta^*(R_{max}, \delta)$, the parameter values for R_{max} and δ must be set. Following Oster's recommendation, I set $R_{max} = \min\{1.3\tilde{R},\ 1\}$ and $\delta = 1$. Consequently, the bounding set for β, which includes the true estimate, is defined as $[\tilde{\beta}, \beta^*(R_{max}, \delta)]$.

Table 6 reports the bound estimates. As Columns (2) and (3) show, in all specifications except when using controls C4 or C5 in *Suskill* regressions, the bounding set excludes zero; that is, the positive coefficients of *EntrLearn* in the baseline LPM regressions are robust to omitted variable bias. As explained in Section 4.2, the much larger coefficients in Column (5) are mainly due to the smaller sample size. Column (1) does not include *EntrOld* (the population ratio of entrepreneurs who have been operating their business for more than 10 years) in the control variables. Thus, I focus on the bounding sets in Columns (2)–(4): the coefficients of *EntrLearn* are in the range of 0.901–2.618 in the *EntrepY5* regressions and 0.771–2.389 in the *TEA* regressions.

Utilizing these bound estimates, I calculate the percentage point changes in *EntrepY5* or *TEA* when *EntrLearn* is increased by 1 SD (see Figure 2). A 1-SD increase in $EntrLearn_{cgae}$ increases the probability that an individual will become an entrepreneur by 0.5–1.5 percentage points in terms of *EntrepY5*, while the mean of *EntrepY5* is 12.0%. Similarly, *TEA*, the mean of which is 10.9%, is increased by 0.4–1.4 percentage points. Figure 2 reports the impact of $Knowent_{it}$ and $EntrOld_{cgaet}$ utilizing the estimated coefficients of the LPM in Table 4. The magnitude of $EntrLearn_{cgae}$'s effects is comparable to that of $EntrOld_{cgaet}$ but much less than that of $Knowent_{it}$. One reason for this smaller effect is a smaller variation (i.e., a smaller SD) in the *cgae*-level average variable, $EntrLearn_{cgae}$. When replacing the SD of $EntrLearn_{cgae}$ (0.006) with the SD calculated from the PIAAC individual-level $EntrLearn_i$ (0.025), the effect becomes 4.4 (= 0.025/0.006) times greater.[24] As discussed in Footnote 20, if I interpret $EntrLearn_{cgae}$ as a proxy for individual i's *own* learning experiences with entrepreneurial tasks as an employee, using $EntrLearn_{cgae}$ generates an attenuation bias resulting from measurement error. Consequently, the effects of individual-level learning experiences would be greater than those of $EntrLearn_i$ in Figure 2.

4.6. Learning Opportunity Environment-Level Estimation

As a final robustness check, I check whether the baseline results are robust at the *cgae*- or *cgaet*-environment level. I collapse all of the individual-level variables used in the previous regressions at either the *cgae* or *cgaet* level. At the *cgae* level, I confirm that an employee's learning opportunities (*EntrLearn*) are positively associated with their self-perceived entrepreneurial skills (*Suskill*) and the entrepreneurship rate (*EntrepY5* or *TEA*) (Figure A1). These positive associations are robust at the *cgaet*-level, even when controlling for various control variables as in the individual-level regressions and instrumenting *EntrLearn* with *Boss* and *BossDistance* (Table A8).

4.7. Implications: The Case of Japan

In this section, I use Japan as an example to discuss the effect of improving employees' opportunities to learn entrepreneurial tasks on the entrepreneurship rate. The entry rate of new firms has been low in Japan, particularly since the collapse of the bubble economy in the 1990s (Honjo 2015). In terms of the GEM data used in the current study,

---

[24] Alternatively, when I perform *cgaet*-level least squares regressions by collapsing all of the variables at the *cgaet* level, the impact gap between *EntrLearn* and *Knowent* shrinks substantially. The corresponding effects become +0.4–0.7 percentage points (*EntrLearn*), +2.7–3.4 percentage points (*Knowent*), and +0.9–1.1 percentage points (*EntrOld*) in terms of *EntrepY5* (see Section 4.6 and Table A8).

the average entrepreneurship rate during 2012–2017 is 6.0% (in terms of *EntrepY5*) or 5.4% (*TEA*) for Japan, both of which is only one-half of the average rate of the 23 countries analyzed in Section 4 and the lowest among them.

Figure A2 shows that the positive relationship between employees' opportunities to learn entrepreneurial tasks and the entrepreneurship rate is also clearly observed at the *cgae*-environment level within Japan. The average $EntrLearn_{cgae}$ of Japan is 0.013, which is less than the 23-county average of 0.015. This is because Japanese employees perform fewer managerial and diversified tasks and more clerical tasks, which leads to a lower $EntrLearn_{cgae}$ (Figure 3). Figure 3 also presents a task set of "BEL041," the *cgae* environment defined by a male Belgian employee, aged 45–54, with a tertiary education and higher. The $EntrLearn_{cgae}$ of BEL041 (0.032) is the third-highest of the 444 *cgae* environments.[25] This is because employees in the BEL041 environment perform more tasks in ICT, management, and finance, more autonomous and diversified tasks, and fewer clerical tasks, all of which result in a high $EntrLearn_{cgae}$. I now assume that the relationship between $EntrLearn_{cgae}$ and the entrepreneurship rate observed in all 23 countries (see Section 4.5) is also applicable to Japan. Then, if the $EntrLearn_{cgae}$ of Japan (0.013) was increased to that of BEL041 (0.032) by achieving the same task set as BEL041, Japan's entrepreneurship rate would increase from the current 6.0% to 7.7%–10.8% in terms of *EntrepY5*. Similarly, the TEA would increase from 5.4% to 6.8%–9.7%. Consequently, Japan's entrepreneurship ranking would increase from the bottom 23rd to between the 12th and the 20th.

5. Results on Entrepreneurial Performance

For the dependent variable representing entrepreneurial performance, I use (i) three binary employment-level indicators ($Emp5_{icgaet}$, $Emp10_{icgaet}$, and $Emp20_{icgaet}$) and (ii) an innovativeness index ($Innov_{icgaet}$), where *i* is an entrepreneur in terms of either

[25] Figure A3 presents similar radar charts for G7 countries other than Japan.

$EntrepY5_{icgaet} = 1$ or $TEA_{icgaet} = 1$. More direct performance measures such as profits or productivity are not available from the GEM data. $EmpZ_{icgaet}$ (where *Z* is either 5, 10, or 20) takes a value of 1 if the "employment," which is defined as the number of people currently working for *i*'s business (including *i* and exclusive subcontractors), is *Z* or more.[26] $Innov_{icgaet}$ measures how innovative *i*'s business is. It is computed as the average of the standardized answers to the following three GEM questions: (i) "*Will all, some, or none of your potential customers consider this product or service new and unfamiliar?*" (ii) "*Right now, are there many, few, or no other businesses offering the same products or services to your potential customers?*" and (iii) "*How long have the technologies or procedures required for this product or service been available? Less than a year, between 1 and 5 years, or longer than 5 years?*"[27]

Because entrepreneurial performance is observed only for entrepreneurs, I estimate a sample selection model using a maximum likelihood estimator (MLE) to control for selection into entrepreneurship (Green 2018: 953; Van de Ven and Van Pragg 1981). I use *Fearfail*, which indicates that the fear of failure prevents *i* from starting a business, as an exclusion restriction assuming that *Fearfail* affects *i*'s decision to become an entrepreneur (which has already been confirmed in Table 4) but does not influence entrepreneurial performance directly.

I first estimate linear equations for all of the performance variables based on the observations with *EntrepY5* = 1. Regarding binary employment-level indicators, I also estimate probit equations (Van de Ven and Van Pragg 1981). Table 7 reports the results. *Rho* indicates the estimated correlation between the error term from the regression equation (or the latent equation, in the probit model) and the error term from the selection equation. If *Rho* is significantly different from zero, the existence of sample selection bias is suspected.

---

[26] Note that the employment level of entrepreneurs' businesses is usually small: the median and the 90th percentile employment level of entrepreneurs' businesses when $EntrepY5_{icgaet} = 1$ is 2 and 10 persons, respectively (based on the regression samples in Column (2) of Table 7).

[27] The answer to each question takes three values: 1 = all, 2 = some, and 3 = none in (i); 1 = many, 2 = few, and 3 = no in (ii); and 1 = less than a year, 2 = 1–5 years, and 3 = longer than 5 years in (iii). The scores in (i) and (iii) are reversed.

Table 7 shows that sample selection bias is likely in Columns (1)–(3), which are based on a large sample size. The coefficients of *EntrLearn* are insignificant when *Emp5* or *Innov* are used as the dependent variables. *EntrLearn* tends to increase *Emp10* and *Emp20*; however, the association is not always statistically significant. These insignificant or weakly positive effects can be interpreted as the size of *positive* skill improvement effect being slightly greater than or the same as that of *negative* entry cost reduction effect (see Section 2.1 for the theoretical background).

When using observations with *TEA* = 1, the results are less robust (Table A9). Similar trends to those in Table 7 are observed when alternative *EntrLearn* indices are used (Table A10). To deal with possible endogeneity bias, I instrument *EntrLearn* with the same instruments (*Boss* and *BossDistance*) used in the occupational choice regressions. To consider both the endogeneity of *EntrLearn* and sample selection bias, I employ an extended regression model (Stata Press 2021). This approach uses MLE; that is, it maximizes the natural logarithm of the joint density of all endogenous variables (outcome variable, selection indicator [i.e., entrepreneur indicator], and possibly endogenous *EnrtLearn*), which is a product of the marginal and conditional distributions. The estimation results in Table A11 show that the exogeneity of *EntrLearn* is not rejected in most cases. Thus, I primarily treat *EntrLearn* as exogenous as in Table 7.

6. Conclusions

Most entrepreneurs accumulate entrepreneurial skills while working as employees. Given the importance of an employee's learning environment, this study has examined the effects of employees' opportunities to learn entrepreneurial tasks on their entrepreneurial activity. It contributes to the literature by directly measuring individual learning opportunities by identifying entrepreneurial tasks and calculating the similarity between employees' tasks and those of entrepreneurs.

Using information on 47 tasks from 31 countries, I first showed that entrepreneurs perform more autonomous and diverse tasks, finance and managerial tasks, and fewer clerical tasks than employees. Employees are considered to have more opportunities to learn entrepreneurial tasks when they experience tasks that are more similar to those of entrepreneurs. Next, using individual-level data from 23 countries in 2012–2017, I found that individuals working in an environment in which employees have more opportunities to learn entrepreneurial tasks have more self-perceived entrepreneurial skills and are more likely to become entrepreneurs. This relationship is generally robust to alternative learning opportunity indices, IV estimation, omitted variable bias, and learning-environment-level regressions. My bound estimates imply that a 1-SD increase in an employee's opportunities to learn entrepreneurial tasks increases the probability that an individual will become an entrepreneur by 0.5–1.5 percentage points (using the *EntrepY5* definition of an entrepreneur), which is 4%–12% of the mean entrepreneurship rate.

More learning opportunities also tend to increase the employment level of an entrepreneur's business, although the results are not very robust. Furthermore, they are not associated with the innovativeness of the business. These weak or insignificant associations with entrepreneurial performance can be interpreted as follows: the negative entry cost reduction effect (which causes lower-ability persons to become entrepreneurs) offsets the positive skill improvement effect.

This study demonstrates that the tasks of employees, which account for 80%–90% of the total employment in advanced countries, are an important determinant of a country's entrepreneurship rate. However, it is not easy to change the task set of employees because it is likely to be influenced by country-specific industry composition, work organization, work style by gender, and age composition. For instance, in a country with an aging population and more gender inequality, young and female employees are promoted more slowly and may therefore have fewer opportunities to learn entrepreneurial tasks. Employees in countries characterized by a consensus decision-making system or a large share of industries requiring coordination among workers may perform more coordination-related tasks. Exploring the factors that determine an employee's task content is left for future research.[28]

References

Aldén, Lina, Mats Hammarstedt, and Emma Neuman. 2017. "All About Balance? A Test of the Jack-of-all-Trades Theory Using Military Enlistment Data." *Labour Economics*, 49: 1-13.

Álvarez, Claudia, David Urbano, and José Ernesto Amorós. 2014. "GEM Research: Achievements and Challenges." *Small Business Economics*, 42: 445–465.

Asuyama, Yoko. 2022. "Doing Boss-like Tasks and Worker Well-being: Job Enrichment Revisited." *Labour*, 36(2): 196–230.

Bhidé, Amar V. 2000. *The Origin and Evolution of New Businesses*. Oxford; New York: Oxford University Press.

Boeri, Tito, Giulia Giupponi, Alan B. Krueger, and Stephen Machin. 2020. "Solo Self-Employment and Alternative Work Arrangements: A Cross-Country Perspective on the Changing Composition of Jobs." *Journal of Economic Perspectives*, 34(1): 170–195.

Burton, M. Diane, Jesper B Sørensen, and Christine M. Beckman. 2002. "7. Coming from Good Stock: Career Histories and New Venture Formation." Lounsbury, M., and Ventresca, M.J. Ed. *Social Structure and Organizations Revisited (Research in the Sociology of Organizations, Vol. 19)*, Bingley, UK: Emerald Group Publishing Limited, pp. 229-262.

Costello, Anna B., and Jason W. Osborne. 2005. "Best Practices in Exploratory Factor

[28] In Table A12, I decompose the variation of employee-level opportunities to learn entrepreneurial tasks ($EntrLearn_i$) by applying Fields's (2003) regression-based variance decomposition technique to the PIAAC data. In the decomposition, the observable variables used in Table 2 are mostly included as predictors for $EntrLearn_i$. The results indicate that around 78% of the variation in $EntrLearn_i$ is explained by residuals; that is, unobservable factors. Thus, the factors that determine $EntrLearn_i$, or an employee's task content, are left for future research using another dataset. Among the observable predictors, occupation (with predictive power of about 10%), industry (4%), boss status (2%–3%), and country (2%–3%) have relatively high predictive power for $EntrLearn_i$. By contrast, gender (0.2%), age and its square (0.2%–0.3%), and education and proficiency in literacy and numeracy (2%) have relatively low predictive power.

Analysis: Four Recommendations for Getting the Most from Your Analysis." *Practical Assessment, Research & Evaluation*, 10(7): 1-9.

DiStefano, Christine, Min Zhu, and Diana Mîndrilă. 2009. "Understanding and Using Factor Scores: Considerations for the Applied Researcher." *Practical Assessment, Research & Evaluation*, 14(20): 1-11.

European Values Study Foundation. 2011. European Values Study 1981-2008, Longitudinal Data File. GESIS Data Archive, Cologne, Germany, ZA4804 Data File Version 2.0.0 (2011-12-30), doi:10.4232/1.11005.

Fields, Gary, S. 2003. "Accounting for Income Inequality and Its Change: A New Method, with Application to the Distribution of Earnings in the United States." *Research in Labor Economics*, 22: 1–38.

GEM (Global Entrepreneurship Monitor) 2022. *Global Entrepreneurship Monitor 2021/2022 Global Report: Opportunity Amid Disruption*. London: GEM.

Giannetti, Mariassunta, Andrei Simonov. 2009. "Social Interactions and Entrepreneurial Activity." *Journal of Economics & Management Strategy*, 18(3): 665–709.

Green, William. H. 2018. *Econometric Analysis*, 8th edition. New York: Pearson.

Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi. 2021. "Learning Entrepreneurship from Other Entrepreneurs?" *Journal of Labor Economics*, 39(1): 135-191.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2006. "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives*, 20(2): 23–48.

Guiso, Luigi, and Fabiano Schivardi. 2011. "What Determines Entrepreneurial Clusters?" *Journal of the European Economic Association*, 9(1): 61–86.

Hartog, Joop, Mirjam van Praag, and Justin van der Sluis. 2010. "If You Are So Smart, Why Aren't You an Entrepreneur? Returns to Cognitive and Social Ability: Entrepreneurs Versus Employees." *Journal of Economics & Management Strategy*, 19(4): 947–989.

Honjo, Yuji. 2015. "Why Are Entrepreneurship Levels So Low in Japan?" *Japan and the World Economy*, 36: 88–101.

Lazear, Edward, P. 2004. "Balanced Skills and Entrepreneurship." *AEA Papers and Proceedings*, 94(2): 208–211.

Lazear, Edward, P. 2005. “Entrepreneurship.” *Journal of Labor Economics*, 23(4): 649–680.

Levine, Ross, and Yona Rubinstein. 2017. “Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?” *Quarterly Journal of Economics*, 132(2): 963–1018.

Liang, James, Hui Wang, and Edward P. Lazear. 2018. “Demographics and Entrepreneurship.” *Journal of Political Economy*, 126(S1): S140-S196.

Lindquist, Matthew J., Joeri Sol, and Mirjam Van Praag. 2015. “Why Do Entrepreneurial Parents Have Entrepreneurial Children?” *Journal of Labor Economics*, 33(2): 269–296.

Lucas, Robert E. Jr. 1978. “On the Size Distribution of Business Firms.” *Bell Journal of Economics*, 9(2): 508–523.

Marino, Marianna, Pierpaolo Parrotta, and Dario Pozzoli. 2012. “Does Labor Diversity Promote Entrepreneurship?” *Economics Letters*, 116(1): 15–19.

Nanda, Ramana, and Jesper B. Sørensen. 2010. “Workplace Peers and Entrepreneurship.” *Management Science*, 56(7): 1116–1126.

OECD (Organisation for Economic Co-operation and Development). 2016. *Skills Matter: Further Results from the Survey of Adult Skills*. Paris: OECD Publishing.

OECD (Organisation for Economic Co-operation and Development), and EU (European Union). 2019. *The Missing Entrepreneurs 2019: Policies for Inclusive Entrepreneurship*. Paris: OECD Publishing.

Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics*, 37(2): 187-204.

Parker, Simon C. 2018. *The Economics of Entrepreneurship, Second Edition*. Cambridge, UK, and New York, US: Cambridge University Press.

Rammstedt, Beatrice, Silke Martin, Anouk Zabal, Ingo Konradt, Débora Maehler, Coviello Anja Perry, Natascha Massing, Daniela Ackermann-Piek, and Susanne Helmschrott. 2016. Programme for the International Assessment of Adult Competencies (PIAAC), Germany - Reduced Version. GESIS Data Archive, Cologne. ZA5845 Data file Version 2.2.0, doi:10.4232/1.12660.

Reynolds, Paul D. 2021. Global Entrepreneurship Monitor [GEM]: Adult Population Survey Data Set, 1998-2017. Inter-university Consortium for Political and Social Research [distributor], 2021-04-12. https://doi.org/10.3886/ICPSR20320.v5

Reynolds, Paul, Niels Bosma, Erkko Autio, Steve Hunt, Natalie De Bono, Isabel Servais, Paloma Lopez-Garcia, and Nancy Chin. 2005. "Global Entrepreneurship Monitor: Data Collection Design and Implementation: 1998-2003." *Small Business Economics*, 24: 205-231.

Sørensen, Jesper B., and Magali A. Fassiotto. 2011. "Organizations as Fonts of Entrepreneurship." *Organization Science*, 22(5): 1322-1331.

Spanjer, Anne, and Arjen van Witteloostuijn. 2017. "The Entrepreneur's Experiential Diversity and Entrepreneurial Performance." *Small Business Economics*, 49, 141–161.

Stata Press. 2021. *Extended Regression Models Reference Manual: Release 17*. College Station, TX: Stata Press.

Stock James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In Donald W. K. Andrews. *Identification and Inference for Econometric Models*. New York: Cambridge University Press: 80–108.

Tåg, Joacim, Thomas Åstebro, and Peter Thompson. 2016. "Hierarchies and Entrepreneurship." *European Economic Review*, 89: 129-147.

Unger, Jens. M., Andreas Rauch, Michael Frese, and Nina Rosenbusch. 2011. "Human Capital and Entrepreneurial Success: A Meta-analytical Review." *Journal of Business Venturing*, 26(3), 341–358.

Van de Ven, Wynand P.M.M., and Bernard M.S. Van Pragg. 1981. "The Demand for Deductibles in Private health Insurance: A Probit Model with Sample Selection." *Journal of Econometrics*, 17(2): 229-252.

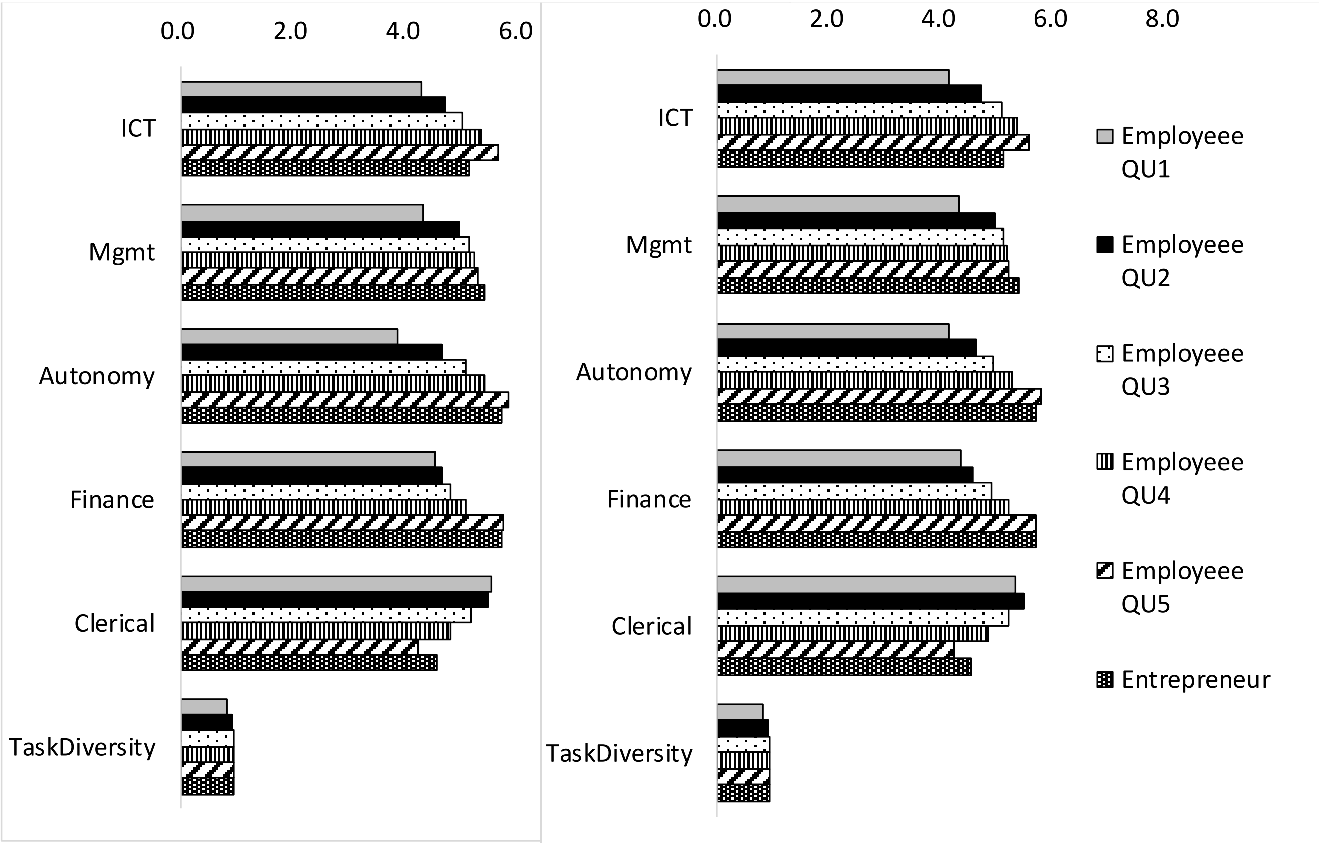
Wagner, Joachim. 2006. "Are Nascent Entrepreneurs 'Jacks-of-all-Trades'? A Test of Lazear's Theory of Entrepreneurship with German Data." *Applied Economics*, 38(20): 2415–2419.

Wooldridge, Jeffrey, M. 2010. *Econometric Analysis of Cross Section and Panel Data, Second Edition*. Cambridge, Massachusetts, and London, England: The MIT Press.

World Values Survey Association. 2015. World Value Survey 1981-2014 Longitudinal Aggregate v.20150418, 2015. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: JDSystems Data Archive, Madrid, Spain.

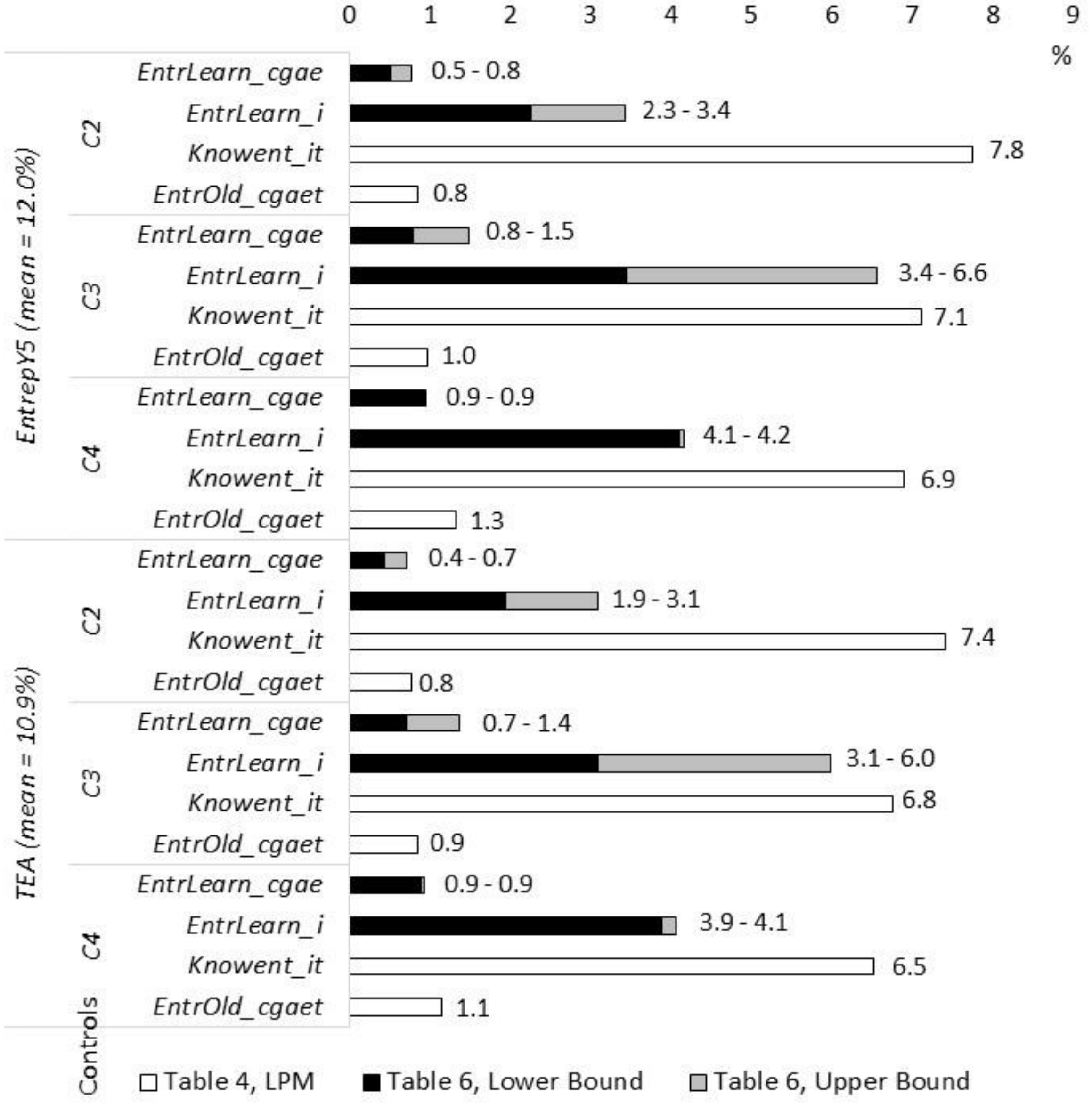
Figure 1. Average Task Scores of Entrepreneurs and Employees by Learning Opportunity Quintile

(a) By $EntrLearn_i$ Quintile (b) By $EntrLearnSQ_i$ Quintile



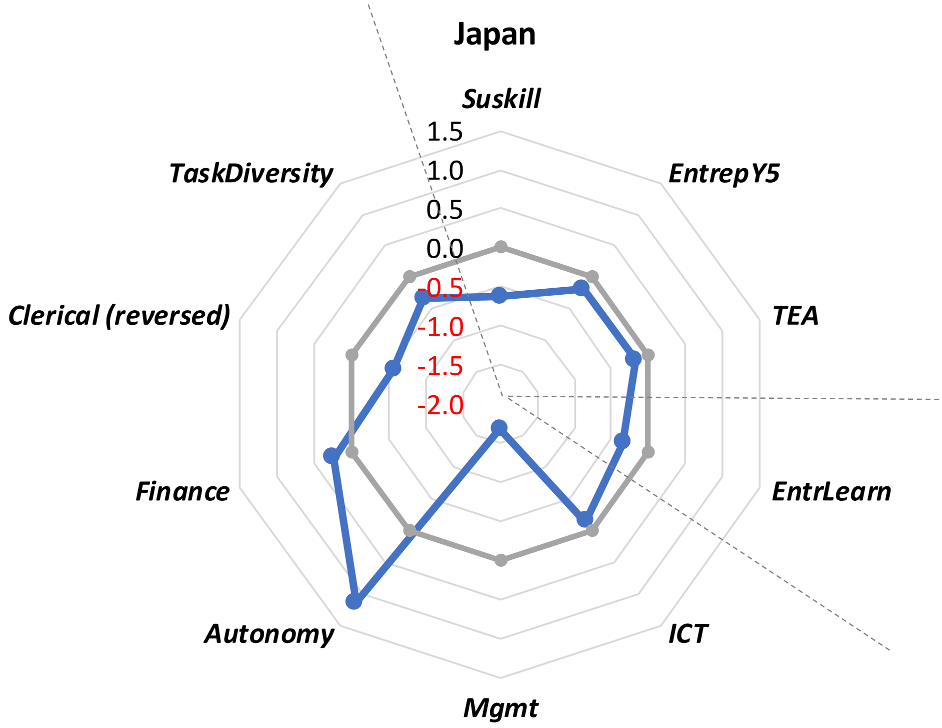
Notes: "QU" represents quintile. The bar graph reports the average of each task variable for entrepreneurs and employees (separated by quintile of "opportunity to learn entrepreneurial tasks"), calculated from PIAAC data. The statistics are weighted by the senate weights in PIAAC (that reflect PIAAC's sampling weights within a country but give each country equal total weight). In (a), learning opportunity, $EntrLearn_i$, is the predicted probability that an employee is an entrepreneur predicted only from the person's five task scores (*ICT*, *Mgmt*, *Autonomy*, *Finance*, and *Clerical*) and task diversity index (*TaskDiversity*). In (b), learning opportunity, $EntrLearnSQ_i$, is predicted from *ICT*, *Mgmt*, *Autonomy*, *Finance*, *Clerical*, and *TaskDiversity* as well as their squared terms. For the construction of these task variables, see Section 3.2. The five task scores are standardized with a mean of 5 and a standard deviation (SD) of 1.

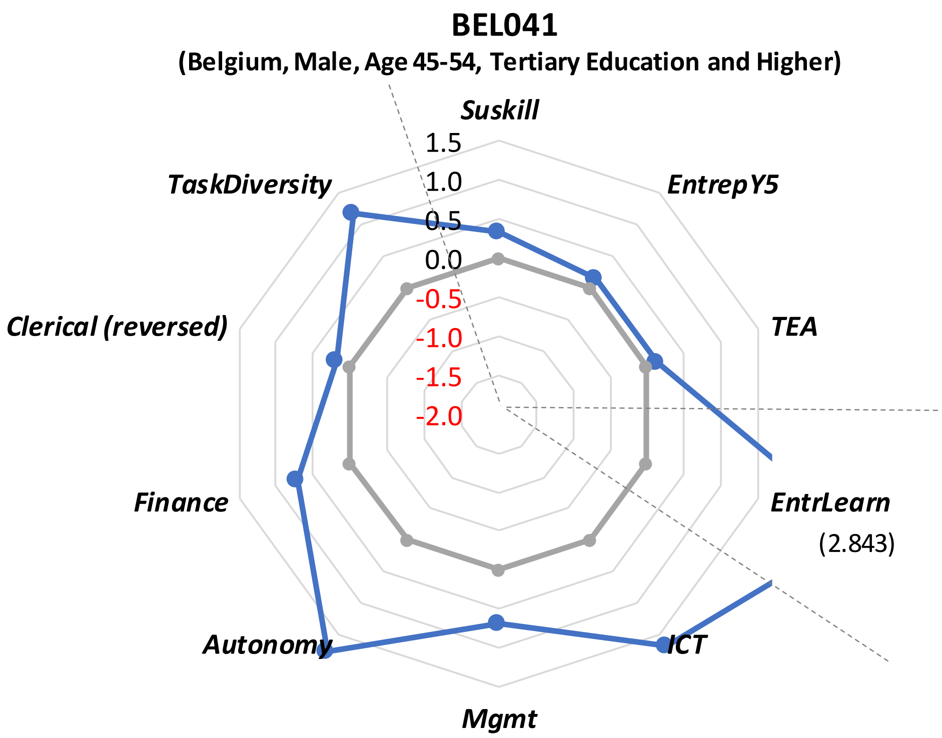
Figure 2. Percentage Point Change in *EntrepY5* or *TEA* when Each Predictor is Increased by One Standard Deviation (SD)



Notes: The impact of *EntrLearn* is based on the lower and upper bounds reported in Table 6 with the corresponding control variables. The impact of $Knowent_{it}$, and $EntrOld_{cgaet}$ is based on the estimated coefficients of the LPM reported in Table 4. The mean of *EntrepY5* and *TEA* and the SD of $EntrLearn_{cgae}$, $Knowent_{it}$, and $EntrOld_{cgaet}$ is based on the values reported in Table 3. The impact of $EntrLearn_i$ utilizes the SD calculated from the individual-level $EntrLearn_i$ based on the 23 countries included in Table 3.

Figure 3. Entrepreneurship Rate, Employees' Learning Opportunities, and Task Scores: Japan and BEL041 (Belgium, Male, Age 45–54, Tertiary Education and Higher)





Notes: Each score is standardized so that the average and standard deviation of all 23 countries is 0 and 1, respectively. The gray line indicates the average score of the 23 countries. The *Clerical* score is reversed because a higher *Clerical* score results in a lower *EntrLearn* (see Table 2). *EntrLearn* and the task variables of Belgium are based on Flanders only. The samples are similar to those in Table 3, but include observations with missing control variables.

Table 1. Tasks and Basic Characteristics of Entrepreneurs and Employees (PIAAC data)

	Total		(a) Entrepreneur	(b) Employee	(a)–(b)
Variable	Mean	Std. dev.	Mean	Mean	Difference
Y = Entrepreneur	0.061	0.239	1.000	0.000	1.000
ICT	5.122	0.984	5.292	5.111	0.182
Mgmt	5.034	0.982	5.415	5.009	0.406
Autonomy	5.049	0.975	5.708	5.006	0.702
Finance	5.017	1.020	5.726	4.971	0.754
Clerical	4.986	1.015	4.546	5.015	−0.469
TaskDiversity	0.923	0.071	0.948	0.922	0.026
TaskDiversityF	0.618	0.158	0.669	0.615	0.055
Female	0.475	0.499	0.268	0.488	−0.220
Age	39.706	11.856	45.007	39.364	5.643
Exp	17.822	11.782	22.935	17.492	5.443
Eduy	13.401	2.773	13.359	13.404	−0.044

Notes: Based on PIAAC regression sample from Column (1) of Table 2. The number of observations is 75,883. Statistics are weighted by the senate weights in PIAAC. For a description of the variables, see Section 3.2 and Table A4.

Table 2. Tasks and Probability of Being an Entrepreneur: Probit Regression Results

	Estimated coefficient				Average marginal effects (AME)			
	(1) Main	(2) Fewer controls	(3) *Task DiveristyF*	(4) Squared terms	(1) Main	(2) Fewer controls	(3) *Task DiveristyF*	(4) Squared terms
ICT	0.082***	0.099***	0.041*	1.179***	0.006***	0.008***	0.003*	-0.004**
	(0.027)	(0.018)	(0.021)	(0.188)	(0.002)	(0.001)	(0.002)	(0.002)
ICT^2				-0.116***				
				(0.018)				
Mgmt	0.087***	0.142***	0.122***	-0.754***	0.006***	0.011***	0.009***	0.005***
	(0.020)	(0.016)	(0.014)	(0.102)	(0.001)	(0.001)	(0.001)	(0.001)
Mgmt^2				0.077***				
				(0.010)				
Autonomy	0.346***	0.357***	0.349***	-0.945***	0.026***	0.027***	0.026***	0.030***
	(0.015)	(0.013)	(0.015)	(0.109)	(0.001)	(0.001)	(0.001)	(0.001)
Autonomy^2				0.122***				
				(0.010)				
Finance	0.191***	0.265***	0.176***	0.552***	0.014***	0.020***	0.013***	0.011***
	(0.016)	(0.012)	(0.014)	(0.125)	(0.001)	(0.001)	(0.001)	(0.001)
Finance^2				-0.036***				
				(0.011)				
Clerical	-0.334***	-0.289***	-0.357***	-0.745***	-0.025***	-0.022***	-0.026***	-0.028***
	(0.014)	(0.011)	(0.014)	(0.070)	(0.001)	(0.001)	(0.001)	(0.001)
Clerical^2				0.039***				
				(0.007)				
TaskDiversity	3.361***	2.798***		-6.698***	0.250***	0.215***		0.531***
	(1.054)	(0.577)		(1.159)	(0.078)	(0.044)		(0.075)
TaskDiversity^2				7.409***				
				(1.106)				
TaskDiversityF			1.681***				0.125***	
			(0.182)				(0.013)	
Controls	C2'	C1'	C2'	C2'	C2'	C1'	C2'	C2'
Observations	75,883	111,676	75,884	75,883	75,883	111,676	75,884	75,883
Pseudo R-squared	0.403	0.387	0.406	0.416				

Notes: Based on entrepreneur and employee PIAAC observations. The binary dependent variable takes the value of 1 if the respondent is an entrepreneur and 0 if the respondent is an employee. All estimations are weighted by the senate weights in PIAAC. Controls C1' include a female dummy, age and its square, years of education, literacy and numeric proficiency scores, learning attitude index, household size, living with a partner dummy, having children dummy, foreign-born dummy, dummies for the public sector and non-profit organization sector, occupation dummies, industry dummies, and country dummies. Controls C2' further include years of work experience and its square, mother's education, father's education, number of books at home, three factor scores for tasks performed outside of work, health status, foreign native language dummy, indices for trust and altruism, and working hours. Robust standard errors are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 3. Summary Statistics of Entrepreneurship Regressions on Learning Opportunities

Variable	Mean	Std. dev.	Variable	Mean	Std. dev.
Suskill	0.448	0.497	*Trust (*)*	0.395	0.190
EntrepY5	0.120	0.325	*LOC (*)*	6.967	0.570
TEA	0.109	0.312	*Creativity (*)*	4.157	0.432
Emp5	0.234	0.423	*N-Ach (*)*	3.583	0.586
Emp10	0.103	0.303	*RiskPref (*)*	3.131	0.569
Emp20	0.048	0.214	*Y2012* (reference)	0.179	0.383
Innov	−0.076	0.684	*Y2013*	0.182	0.386
EntrLearn	0.015	0.006	*Y2014*	0.178	0.383
Female	0.477	0.499	*Y2015*	0.147	0.354
Age	40.406	11.897	*Y2016*	0.168	0.374
Edu1	0.009	0.094	*Y2017*	0.146	0.353
Edu2	0.018	0.134	Austria	0.026	0.159
Edu3	0.107	0.309	Belgium	0.038	0.191
Edu4 (reference)	0.356	0.479	Canada	0.048	0.214
Edu5	0.162	0.369	Cyprus	0.019	0.136
Edu6	0.325	0.468	Czech Republic	0.010	0.101
Edu7	0.023	0.150	Denmark	0.019	0.137
HHsize	3.058	1.494	Estonia	0.058	0.234
Knowent	0.356	0.479	Finland	0.049	0.216
EntrOld	3.800	3.891	France	0.046	0.210
Fearfail	0.471	0.499	Germany	0.056	0.230
LearnAttitude	0.052	0.322	Ireland	0.056	0.230
Htask_ICT	0.042	0.413	Italy	0.041	0.198
Htask_Math	−0.181	0.234	Japan	0.035	0.184
Htask_Clerical	0.017	0.252	Netherlands	0.063	0.244
Fulltime (*)	0.830	0.173	Norway	0.045	0.207
Permanent (*)	0.756	0.162	Poland	0.048	0.214
Estsize (*)	192.007	71.375	Russia	0.033	0.179
Tenure (*)	8.879	5.281	Slovak Republic	0.052	0.222
Opport (*)	0.392	0.488	South Korea	0.039	0.194
Nbgoodc (*)	0.551	0.497	Spain	0.042	0.201
Nbstatus (*)	0.674	0.469	Sweden	0.056	0.230
Nbmedia (*)	0.548	0.498	UK (reference)	0.063	0.242
			US	0.056	0.230

Notes: The statistics of the dependent variables are based on the corresponding Column (2) in Table 4 or Column (2) in Table 7. Those of the other variables are based on Column (2) of the *EntrepY5* regression in Table 4 (237,074 observations), except for variables with an asterisk (*), which are based on Column (5) (113,096 observations). For a description of the variables, see Table A5. The statistics are weighted by the senate weights in GEM (that reflect the GEM's sampling weights within a country but give each country equal total weight).

Table 4. Employees' Learning Opportunities, Skill Acquisition, and Occupational Choice: Baseline Results

Dep. Var.	*Suskill*					*EntrepY5*					*TEA*				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
LPM model															
EntrLearn	3.165***	2.452***	2.882***	2.178***	3.567***	1.327***	0.901**	1.374***	1.660***	3.138***	1.140***	0.771*	1.234***	1.554***	3.056***
	(0.615)	(0.676)	(0.735)	(0.819)	(1.115)	(0.405)	(0.432)	(0.474)	(0.581)	(0.741)	(0.395)	(0.418)	(0.454)	(0.561)	(0.736)
Knowent	0.227***	0.227***	0.225***	0.204***	0.214***	0.162***	0.162***	0.149***	0.144***	0.149***	0.155***	0.155***	0.141***	0.136***	0.141***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
EntrOld		0.005***	0.006***	0.006***	0.006***		0.002***	0.002***	0.003***	0.004***		0.002***	0.002***	0.003***	0.004***
		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
Fearfail	-0.135***	-0.135***	-0.129***	-0.122***	-0.125***	-0.081***	-0.082***	-0.080***	-0.081***	-0.085***	-0.074***	-0.074***	-0.072***	-0.073***	-0.075***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
R-squared	0.151	0.152	0.147	0.155	0.165	0.102	0.102	0.094	0.107	0.114	0.099	0.099	0.090	0.103	0.109
Probit model (Average marginal effects [AME])															
EntrLearn	3.374***	2.825***	3.241***	2.555***	3.845***	1.176***	0.762*	1.068**	1.287**	2.557***	0.990***	0.638*	0.930**	1.168**	2.401***
	(0.608)	(0.660)	(0.718)	(0.796)	(1.111)	(0.354)	(0.392)	(0.417)	(0.509)	(0.693)	(0.339)	(0.368)	(0.386)	(0.476)	(0.665)
Knowent	0.226***	0.226***	0.224***	0.202***	0.212***	0.157***	0.157***	0.144***	0.137***	0.143***	0.150***	0.150***	0.136***	0.129***	0.135***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
EntrOld		0.005***	0.005***	0.006***	0.006***		0.002***	0.002***	0.003***	0.004***		0.002***	0.002***	0.003***	0.003***
		(0.001)	(0.001)	(0.001)	(0.001)		0.000	0.000	(0.001)	(0.001)		0.000	0.000	(0.001)	(0.001)
Fearfail	-0.134***	-0.134***	-0.128***	-0.121***	-0.124***	-0.079***	-0.079***	-0.077***	-0.077***	-0.081***	-0.072***	-0.072***	-0.070***	-0.069***	-0.072***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Pseudo R-squared	0.119	0.119	0.115	0.122	0.131	0.138	0.138	0.134	0.145	0.151	0.142	0.142	0.137	0.147	0.155
Controls	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5
Observations	252,188	251,544	224,306	152,228	122,846	237,709	237,074	210,895	140,520	113,096	234,860	234,229	208,389	138,660	111,526

Notes: The binary dependent variable is either *Suskill* (an indicator for sufficient entrepreneurial skills and knowledge acquired), *EntrepY5* (an indicator for an entrepreneur whose business is less than 5 years old), or *TEA* (an entrepreneur whose business is less than 3.5 years old). For more details, see Section 4.2. *EntrLearn* is the average employee's opportunities to learn entrepreneurial tasks in the *cgae* environment (country * gender * age group * education group). *Knowent* is a dummy for knowing "*someone personally who started a business in the past 2 years.*" *EntrOld* is the *cgaet*-level population ratio (%) of old entrepreneurs (owner-managers of a new firm founded more than 10 years before year *t*). *Fearfail* is a dummy for agreeing that "*fear of failure would prevent you from starting a business.*" Controls C1 include *icgaet*-level covariates: a female dummy, age and its square, education level dummies, household size, year dummies, and country dummies. C2 includes C1 plus the following *cgae*-level variables: average employee's learning attitude and three factor scores for tasks performed outside of work. C3 includes C2 and the *cgae*-level covariates including the ratio of full-time employees (*Fulltime*), the ratio of employees with indefinite employment contracts (*Permanent*), the average number of employees at the workplace (*Estsize*), and the average tenure of workers (*Tenure*). C4 includes C3 plus the *icgaet*-level variables including perceived local opportunities for starting a business (*Opport*), perceived attitude of citizens toward the desirability and status of an entrepreneur's job (*Nbgoodc* and *Nbstatus*), and the frequency of media coverage of entrepreneurial success (*Nbmedia*). C5 includes C4 plus *cgae*-level average worker psychological traits such as trust, locus of control, focus on creativity, need for achievement, and risk preference. All estimations are weighted by the senate weights in GEM. Standard errors clustered by *cgae* environment are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 5. Employees' Learning Opportunities, Skill Acquisition, and Occupational Choice: Instrumental Variable (IV) Estimation Results

Dep. Var.	*Suskill*					*EntrepY5*					*TEA*				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
(a) 2SLS LPM model															
EntrLearn	5.125***	2.951	4.721**	5.114*	7.423***	4.116***	3.849***	5.932***	6.407***	6.025***	3.493***	3.132***	4.929***	5.681***	5.107***
	(1.611)	(2.131)	(2.370)	(2.641)	(2.674)	(0.895)	(1.231)	(1.612)	(1.910)	(1.843)	(0.843)	(1.168)	(1.479)	(1.730)	(1.657)
R-squared	0.150	0.151	0.145	0.154	0.165	0.101	0.101	0.092	0.105	0.113	0.098	0.098	0.089	0.100	0.109
Endogeneity test	1.863	0.089	0.755	1.619	2.695	12.341***	6.380**	9.932***	7.582***	2.873*	10.117***	4.567**	7.632***	6.774***	1.712
Overidentification test	0.076	0.103	0.147	0.239	0.150	0.572	0.877	0.123	0.072	0.018	0.737	1.188	0.384	0.132	0.020
First-stage (Y = EntrLearn)															
Boss	0.018***	0.014***	0.015***	0.014***	0.015***	0.018***	0.014***	0.015***	0.013***	0.015***	0.018***	0.014***	0.015***	0.013***	0.015***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
BossDistance	-0.008***	-0.006***	-0.005**	-0.005**	-0.006**	-0.008***	-0.006***	-0.006**	-0.005**	-0.006**	-0.008***	-0.006***	-0.006**	-0.005**	-0.006**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
1 st Stage F stat.	67.210	33.263	23.928	20.990	19.513	65.888	34.720	23.586	20.321	18.744	65.883	34.773	23.594	20.348	18.787
(b) Two stage probit model by control function approach															
Average marginal effects (AME)															
EntrLearn	5.406***	3.444***	5.061***	5.613***	7.655***	3.493***	3.496***	5.178***	5.668***	5.513***	3.001***	2.953***	4.328***	5.088***	4.570***
	(0.875)	(1.214)	(1.385)	(1.753)	(2.010)	(0.556)	(0.840)	(0.928)	(1.244)	(1.302)	(0.563)	(0.798)	(0.946)	(1.231)	(1.266)
Coefficients of 1st stage residual															
1st stage residual	-7.251	-2.238	-5.927	-10.199	-13.080*	-16.716***	-17.400***	-26.969***	-27.317***	-19.463*	-15.497***	-15.815**	-23.990***	-26.411***	-15.462
	(4.888)	(6.089)	(6.700)	(7.348)	(7.250)	(5.384)	(6.466)	(7.873)	(9.327)	(10.022)	(5.271)	(6.391)	(7.672)	(8.936)	(9.819)
Pseudo R-squared	0.118	0.118	0.114	0.121	0.131	0.139	0.139	0.135	0.145	0.152	0.142	0.142	0.137	0.147	0.155
Controls	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5
Observations	250,557	249,939	222,829	151,151	122,675	236,326	235,716	209,625	139,607	112,952	233,505	232,899	207,145	137,771	111,386

Notes: In (a), the LPM model is estimated by two-stage least squares (2SLS). In (b), a two-stage probit model is estimated by following the control function approach (Wooldridge 2010: 586–589). Instruments for *EntrLearn* are *Boss* (the ratio of bosses to all employees in a *cgae* cell) and *BossDistance* (average expected distance to become a boss for non-boss employees in a *cgae* cell). For more details, see Section 4.4. All estimations are weighted by the senate weights in GEM. Standard errors clustered by *cgae* environment are in parentheses except for those of AME, which are computed by bootstrapping the entire two-step procedure 250 times. Like in Table 4, *Knowent* and *Fearfail* are controlled in all columns, and *EntrOld* is controlled in Columns (2)–(5). For the other control variables, see Table 4. The endogeneity test is the chi-squared statistic, which tests the exogeneity of *EntrLearn* (the null hypothesis). The overidentification test is the chi-squared statistic, which tests whether the IVs are valid (i.e., uncorrelated with the error term) (null hypothesis). The first-stage *F* statistic is the Kleibergen–Paap–Wald rk *F* statistic. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 6. Employees' Learning Opportunities, Skill Acquisition, and Occupational Choice: Bound Estimates for the LPM Model

Specification	(1) Controlled effect $\tilde{\beta}$ (Std. error)		(2) Bounding set $[\tilde{\beta}, \beta^{*}(\min\{1.3\tilde{R}, 1\}, \delta)$, with $\delta = 1$	(3) (2) excludes 0?	(4) δ for $\beta = 0$ given R_{max}	(5) R_{max}
Y = Suskill						
Controls C1	3.165***	(0.615)	[3.165, 3.723]	Yes	2.527	0.197
Controls C2	2.452***	(0.677)	[2.452, 1.892]	Yes	1.477	0.198
Controls C3	2.882***	(0.736)	[2.882, 3.397]	Yes	2.045	0.191
Controls C4	2.178***	(0.820)	[2.178, -2.604]	No	0.645	0.202
Controls C5	3.567***	(1.118)	[3.567, -3.289]	No	0.827	0.215
Y = EntrepY5						
Controls C1	1.327***	(0.405)	[1.327, 2.273]	Yes	8.365	0.132
Controls C2	0.901**	(0.432)	[0.901, 1.365]	Yes	2.713	0.133
Controls C3	1.374***	(0.474)	[1.374, 2.618]	Yes	5.636	0.123
Controls C4	1.660***	(0.582)	[1.660, 1.638]	Yes	1.790	0.139
Controls C5	3.138***	(0.743)	[3.138, 8.719]	Yes	3.209	0.148
Y = TEA						
Controls C1	1.140***	(0.396)	[1.140, 1.988]	Yes	9.744	0.128
Controls C2	0.771*	(0.418)	[0.771, 1.230]	Yes	3.000	0.129
Controls C3	1.234***	(0.454)	[1.234, 2.389]	Yes	6.160	0.118
Controls C4	1.554***	(0.562)	[1.554, 1.618]	Yes	1.876	0.133
Controls C5	3.056***	(0.738)	[3.056, 8.113]	Yes	2.920	0.142

Notes: The results in Column (1) are from the LPM in Table 4 with the corresponding control variables. The results in Columns (2) and (4) are calculated using the Stata code "psacalc" developed by Oster (2019). $R_{max} = \min\{1.3\tilde{R}, 1\} = 1.3\tilde{R}$ in this table. See Section 4.5 for an explanation of bound estimates. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

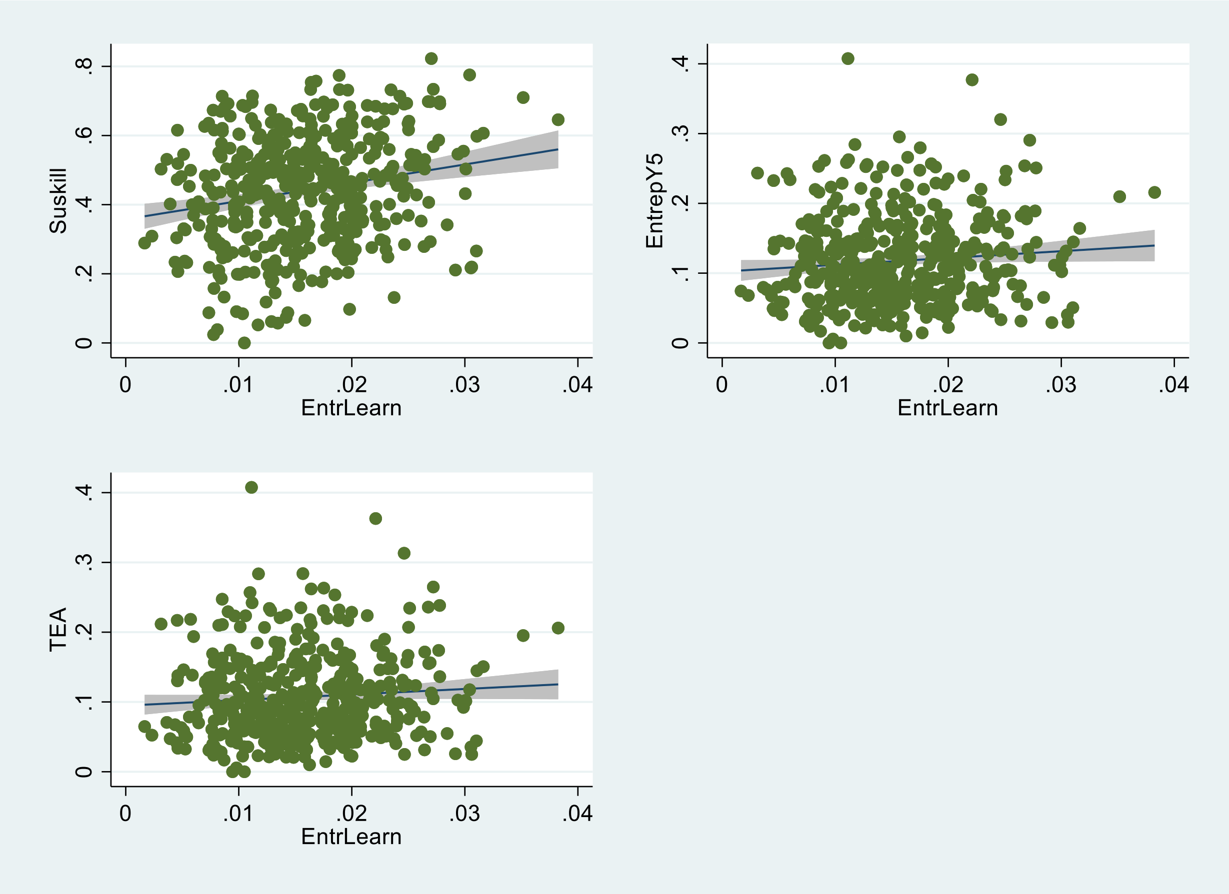
Table 7. Employees' Learning Opportunities and Entrepreneur Performance: Sample Selection Model with Entrepreneur (*EntrepY5* = 1) Definition

	Linear selection model					Probit selection model (coefficients)				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Y = Emp5										
EntrLearn	-0.183	-1.082	-1.214	-0.949	-3.686	-0.444	-3.167	-3.930	-3.190	-12.696
	(1.193)	(1.257)	(1.308)	(1.678)	(2.241)	(4.002)	(4.367)	(4.606)	(5.662)	(7.929)
Rho	-0.044	-0.045	-0.047	-0.080	-0.109	-0.165	-0.157	-0.154	-0.043	-0.004
	(0.044)	(0.045)	(0.047)	(0.080)	(0.109)	(0.080)	(0.080)	(0.088)	(0.116)	(0.131)
Test for *Rho* = 0	6.532**	5.800**	4.950*	0.202	0.000	4.134**	3.693*	2.974*	0.136	0.001
Obs.	227,962	227,352	203,451	135,067	108,915	227,962	227,352	203,451	135,067	108,915
Obs. selected	16,835	16,803	14,599	10,645	8,956	16,835	16,803	14,599	10,645	8,956
Y = Emp10										
EntrLearn	1.743**	1.640*	1.562*	2.640**	0.659	9.397**	8.823	7.689	12.216*	3.569
	(0.778)	(0.897)	(0.874)	(1.143)	(1.351)	(4.603)	(5.403)	(5.425)	(6.507)	(8.760)
Rho	-0.037	-0.038	-0.040	-0.060	-0.078	-0.194	-0.187	-0.202	-0.099	-0.073
	(0.037)	(0.038)	(0.040)	(0.060)	(0.078)	(0.103)	(0.103)	(0.115)	(0.145)	(0.172)
Test for *Rho* = 0	7.546***	6.958***	6.736***	1.112	0.328	3.398*	3.139*	2.925*	0.457	0.179
Obs.	227,962	227,352	203,451	135,067	108,915	227,962	227,352	203,451	135,067	108,915
Obs. selected	16,835	16,803	14,599	10,645	8,956	16,835	16,803	14,599	10,645	8,956
Y = Emp20										
EntrLearn	0.560	1.026*	1.377**	2.331***	0.298	3.222	6.847	10.329	16.714*	-4.056
	(0.523)	(0.619)	(0.624)	(0.862)	(0.934)	(5.479)	(6.567)	(6.977)	(8.804)	(12.021)
Rho	-0.040	-0.040	-0.038	-0.045	-0.056	-0.147	-0.140	-0.179	-0.169	-0.173
	(0.040)	(0.040)	(0.038)	(0.045)	(0.056)	(0.133)	(0.134)	(0.147)	(0.173)	(0.208)
Test for *Rho* = 0	2.720*	2.591	3.656*	2.410	1.190	1.179	1.066	1.416	0.915	0.659
Obs.	227,962	227,352	203,451	135,067	108,915	227,962	227,352	203,451	135,067	108,915
Obs. selected	16,835	16,803	14,599	10,645	8,956	16,835	16,803	14,599	10,645	8,956
Y = Innov										
EntrLearn	-0.208	0.473	0.433	-1.132	-3.284					
	(1.748)	(1.898)	(1.851)	(2.174)	(2.874)					
Rho	-0.082	-0.084	-0.130	-0.050	0.063					
	(0.045)	(0.045)	(0.043)	(0.072)	(0.156)					
Test for *Rho* = 0	3.258*	3.460*	8.899***	0.480	0.163					
Obs.	235,584	234,953	209,107	139,445	112,272					
Obs. selected	24,457	24,404	20,255	15,023	12,313					
Controls	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5

Notes: The dependent variable is either (a) a binary employment-level indicator (*EmpZ*, where *Z* is 5, 10, or 20, which indicates employment of the entrepreneur's business is Z or more), or (b) the innovativeness index of the entrepreneurial business (*Innov*). To account for the selection of entrepreneurship, I use *Fearfail* (indicating that fear of failure prevents the respondent from starting a business) as an exclusion restriction. All of the estimations are weighted by the senate weights in GEM. Standard errors clustered by *cgae* environment are in parentheses. *Knowent* is controlled in all columns, and *EntrOld* is controlled for in Columns (2)–(5). For the other control variables, see Table 4. *Rho* indicates the estimated correlation between the error term from the regression equation (or latent equation in the probit model) and the error term from the selection equation. The test for *Rho* = 0 (which indicates no selection) is the chi-squared statistic. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

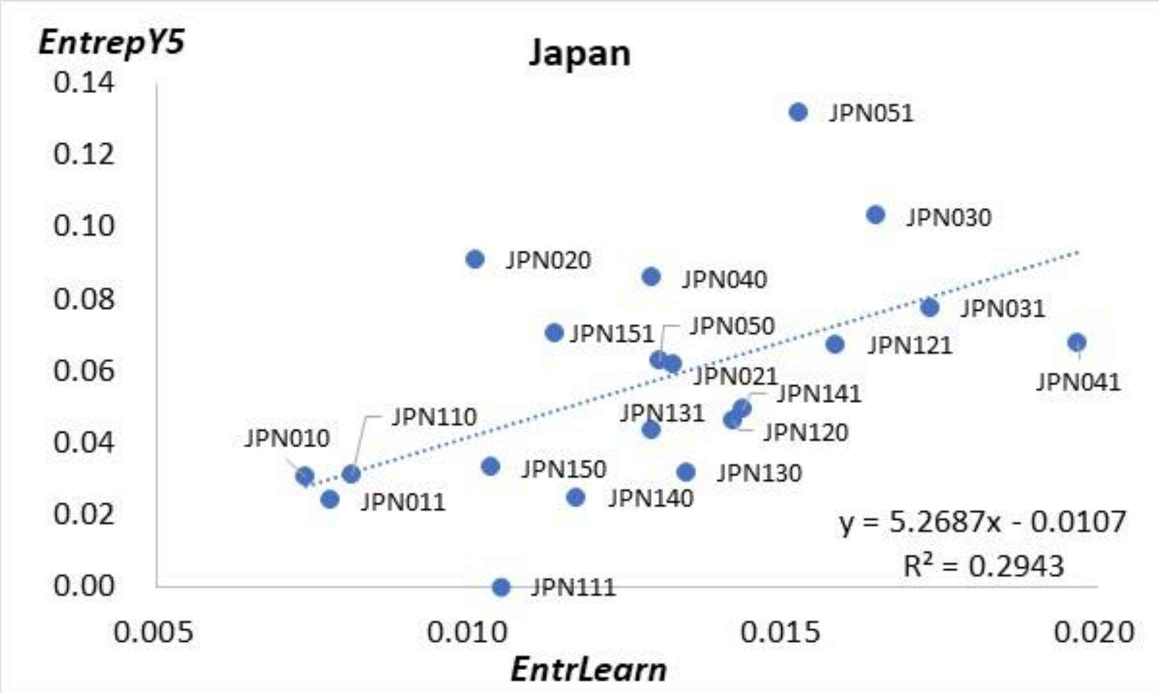
Appendix A

Figure A1. Employees' Learning Opportunities, Skill Acquisition, and Occupational Choice at the *cgae* level



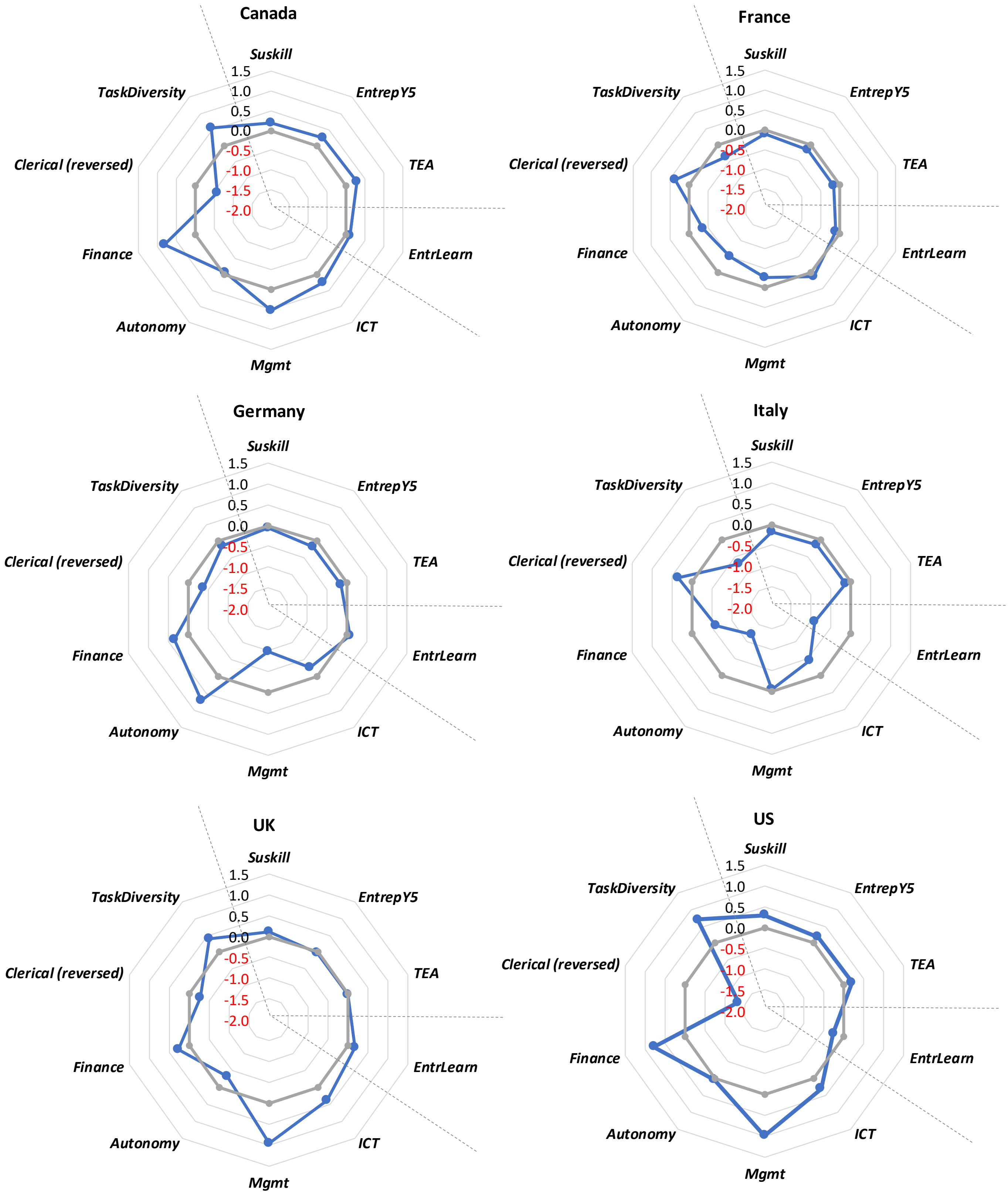
Notes: Each data point plots an average *y*-axis variable (*Suskill*, *EntrepY5*, or *TEA*) and *EntrLearn* (*x*-axis variable) of the *cgae* environment. The GEM's senate weights are applied. The solid line indicates fitted values. The shaded area indicates the 95% confidence interval.

Figure A2. Employees' Learning Opportunities and Entrepreneurship Rate (*EntrepY5*) in Japan at the *cgae* level



Notes: This figure plots the average *EntrLearn* and *EntrepY5* of each *cgae* environment in Japan. The GEM's senate weights are applied. Regarding the data label, the 3-digit number after JPN, respectively, indicates gender (0: male, 1: female), age group (1: age 16–24, 2: 25–34, 3: 35–44, 4: 45–54, 5: 55–65), and education level (0: up to post-secondary non-tertiary education, 1: tertiary education and higher).

Figure A3. Entrepreneurship Rate, Employees' Learning Opportunities, and Task Scores: G7 Countries Other Than Japan



Notes: The same as Figure 3.

Table A1. Description of 47 Tasks at Work

Highest-loading factor	Task	PIAAC Question	Description
ICT	*ICTmail*	G_Q05a	Use e-mail
	ICTinternet	G_Q05c	Use the Internet to better understand work-related issues
	Wmail	G_Q02a	Write letters, memos, or e-mails
	Rmail	G_Q01b	Read letters, memos, or e-mails
	ICTword	G_Q05f	Use a word processor, for example Word
	ICTec	G_Q05d	Conduct transactions on the Internet, for example, buying or selling products or services or banking
	ICTexcel	G_Q05e	Use spreadsheet software, for example Excel
	Rnews	G_Q01c	Read articles in newspapers, magazines, or newsletters
	Rjournal	G_Q01d	Read articles in professional journals or scholarly publications
	Manuwork	F_Q06b	Work physically for a long period of time
	ICTchat	G_Q05h	Participate in real-time discussions on the Internet, for example, in online conferences or chat groups
	Wform	G_Q02d	Fill in forms
	Wnews	G_Q02b	Write articles for newspapers, magazines, or newsletters
Mgmt	*Planother*	F_Q03b	Plan the activities of others
	Persuade	F_Q04a	Persuade or influence people
	Infoshare	F_Q02a	Share work-related information with co-workers
	Planown	F_Q03a	Plan own activities
	Advise	F_Q02e	Advise people
	Teach	F_Q02b	Instruct, train, or teach people individually or in groups
	Negotiate	F_Q04b	Negotiate with people either inside or outside of the organization
	Mnghour	F_Q03c	Organize own time
	Cooperate (*)	F_Q01b	Cooperate or collaborate with co-workers
	PSeasy	F_Q05a	Faced by relatively simple problems that take no more than 5 minutes to find a good solution
	LearnOJT	D_Q13b	Learning-by-doing from the tasks being performed
	LearnCWboss	D_Q13a	Learn new work-related things from co-workers or supervisors
	Infonew	D_Q13c	Keep up to date with new products or services
	PSdiff	F_Q05b	Confronted with more complex problems that take at least 30 minutes to find a good solution
	Fingerwork	F_Q06c	Use skill or accuracy with hands or fingers
Autonomy	*AutoWay (**)*	D_Q11b	Choose or change how to do own work
	AutoOrder (**)	D_Q11a	Choose or change the sequence of own tasks
	AutoSpeed (**)	D_Q11c	Choose or change the speed or rate of own work
	AutoHour (**)	D_Q11d	Choose or change working hours
Finance	*Nprice*	G_Q03b	Calculate prices, costs, or budgets
	Nfraction	G_Q03c	Use or calculate fractions, decimals or percentages
	Ncalculator	G_Q03d	Use a calculator, either hand-held or computer-based
	Nalgebra	G_Q03g	Use simple algebra or formulas
	Rfinstat	G_Q01g	Read bills, invoices, bank statements, or other financial statements

	Sales	F_Q02d	Sell a product or a service
Clerical	*Rmanual*	G_Q01f	Read manuals or reference materials
	Ngraph	G_Q03f	Prepare charts, graphs, or tables
	Rbook	G_Q01e	Read books
	Rgraph	G_Q01h	Read diagrams, maps, or schematics
	Rinstr	G_Q01a	Read directions or instructions
	Nmath	G_Q03h	Use more advanced math or statistics such as calculus, complex algebra, trigonometry, or use of regression techniques
	Wreport	G_Q02c	Write reports
	Present	F_Q02c	Make speeches or give presentations in front of five or more people
	ICTprogram	G_Q05g	Use a programming language to program or write computer code

Note: For all questions except those with (*) and (**), respondents are asked to rate the frequency of each task according to a 5-point scale. The original scale is as follows: 1 = *never*, 2 = *less than once a month*, 3 = *less than once a week but at least once a month*, 4 = *at least once a week but not every day*, and 5 = *every day*. *Cooperate* (*) is measured by the 5-point scale proportion of time usually spent: 1 = *none of the time*, 2 = *up to a quarter of the time*, 3 = *up to half of the time*, 4 = *more than half of the time*, 5 = *all the time*. Four autonomy variables with (**) is measured by the 5-point scale degree: 1 = *not at all*, 2 = *very little*, 3 = *to some extent*, 4 = *to a high extent*, 5 = *to a very high extent*. For more information on the highest-loading factor, which is based on exploratory factor analysis, see Section 3.2.
Source: The original source is a PIAAC background questionnaire. Variable names and descriptions for all of the variables except those marked with (*) and (**) are generally taken from Table A2 of Asuyama (2022) ("Doing Boss-like Tasks and Worker Well-being: Job Enrichment Revisited," Yoko Asuyama, *Labour*, 2022, Volume 36, Issue 2, doi.org/10.1111/labr.12217, ©2022 Fondazione Giacomo Brodolini and John Wiley & Sons Ltd.).

Table A2. Average of 47 Standardized Task Scores of Entrepreneurs and Employees

Highest-loading factor	Task	(1) Entrepreneur	(2) Employee	Difference: (1)–(2)
ICT	*ICTmail*	0.192	0.012	0.179
	ICTinternet	0.225	−0.002	0.227
	Wmail	0.140	0.021	0.119
	Rmail	0.236	0.014	0.221
	ICTword	0.012	0.031	−0.019
	ICTec	0.786	−0.077	0.863
	ICTexcel	0.062	0.032	0.029
	Rnews	0.454	0.035	0.489
	Rjournal	0.337	−0.030	0.368
	Manuwork	0.120	−0.023	0.142
	ICTchat	0.069	−0.006	0.076
	Wform	−0.127	0.053	−0.181
	Wnews	0.057	−0.006	0.063
Mgmt	*Planother*	0.777	−0.017	0.794
	Persuade	0.317	−0.013	0.329
	Infoshare	−0.094	0.113	−0.207
	Planown	0.417	−0.051	0.468
	Advise	0.213	0.007	0.206
	Teach	0.238	0.022	0.216
	Negotiate	0.515	−0.044	0.560
	Mnghour	0.415	−0.058	0.472
	Cooperate	0.052	0.134	−0.082
	PSeasy	0.121	0.014	0.106
	LearnOJT	−0.042	−0.003	−0.039
	LearnCWboss	−0.075	0.102	−0.177
	Infonew	0.141	−0.011	0.152
	PSdiff	0.153	−0.001	0.155
	Fingerwork	0.015	−0.006	0.021
Autonomy	*AutoWay*	0.484	−0.073	0.558
	AutoOrder	0.512	−0.079	0.591
	AutoSpeed	0.437	−0.072	0.509
	AutoHour	0.809	−0.143	0.952
Finance	*Nprice*	0.775	−0.092	0.867
	Nfraction	0.430	−0.029	0.459
	Ncalculator	0.528	−0.032	0.560
	Nalgebra	0.132	0.001	0.132
	Rfinstat	0.844	−0.082	0.926
	Sales	0.808	−0.095	0.903
Clerical	*Rmanual*	0.004	0.010	−0.006
	Ngraph	0.033	0.018	0.015
	Rbook	0.047	−0.015	0.062
	Rgraph	0.084	0.009	0.075
	Rinstr	−0.041	0.035	−0.076

Nmath	0.031	0.002	0.030
Wreport	−0.153	0.046	−0.199
Present	0.004	0.020	−0.016
ICTprogram	−0.018	0.006	−0.024

Notes: Figures indicate the average value of the standardized task scores (standardized *approximated* scores are used except for four autonomy task variables) of entrepreneurs and employees, computed from PIAAC. Cells with an absolute difference greater than 0.3 are highlighted. The statistics are weighted by the senate weights in PIAAC. For more information on the highest-loading factor, which is based on exploratory factor analysis, see Section 3.2. For an explanation of each task item, see Table A1.

Table A3. Factor Loadings of 47 Task Items on the Extracted Factors (Main Analysis)

Factor	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor name	*ICT*	*Mgmt*	*Autonomy*	*Finance*	*Clerical*
Task item					
ICTmail	**0.879**	0.110	0.073	0.042	−0.010
ICTinternet	**0.787**	0.102	0.067	0.064	0.122
Wmail	**0.766**	0.129	0.097	0.125	0.091
Rmail	**0.756**	0.160	0.074	0.077	0.072
ICTword	**0.747**	0.028	0.103	0.129	0.154
ICTec	**0.670**	0.089	0.078	0.178	0.000
ICTexcel	**0.643**	−0.006	0.080	0.235	0.144
Rnews	**0.463**	0.079	0.058	0.022	0.283
Rjournal	**0.442**	0.131	0.040	0.012	0.405
Manuwork	**−0.417**	0.208	−0.116	0.022	−0.049
ICTchat	**0.309**	0.034	0.047	0.025	0.152
Wform	**0.287**	0.106	−0.008	0.269	0.272
Wnews	**0.203**	0.024	0.035	−0.010	0.179
Planother	0.147	**0.549**	0.063	0.198	0.000
Persuade	0.271	**0.519**	0.040	0.245	0.092
Infoshare	0.178	**0.519**	0.045	0.066	0.087
Planown	0.143	**0.510**	0.102	0.098	−0.085
Advise	0.318	**0.505**	0.057	0.217	0.069
Teach	0.160	**0.491**	0.030	0.063	0.215
Negotiate	0.255	**0.488**	0.011	0.326	0.096
Mnghour	0.125	**0.442**	0.148	0.083	−0.090
Cooperate	−0.048	**0.429**	−0.013	0.059	0.068
PSeasy	0.198	**0.417**	0.018	0.178	0.164
LearnOJT	0.107	**0.362**	0.016	−0.054	0.285
LearnCWboss	0.178	**0.343**	0.032	−0.035	0.301
Infonew	0.248	**0.341**	0.053	0.082	0.293
PSdiff	0.234	**0.328**	0.013	0.106	0.302
Fingerwork	−0.123	**0.275**	−0.111	0.072	0.051
AutoWay	0.112	0.055	**0.810**	0.015	0.031
AutoOrder	0.118	0.046	**0.805**	0.027	−0.015
AutoSpeed	0.064	0.023	**0.793**	0.022	−0.006
AutoHour	0.109	−0.028	**0.695**	−0.013	0.021
Nprice	0.142	0.200	−0.002	**0.659**	0.009
Nfraction	0.321	0.141	0.031	**0.599**	0.150
Ncalculator	0.297	0.133	0.043	**0.575**	0.019
Nalgebra	0.196	0.033	0.061	**0.405**	0.222
Rfinstat	0.343	0.179	0.045	**0.397**	−0.009
Sales	0.057	0.328	0.029	**0.361**	−0.106
Rmanual	0.275	0.172	−0.037	0.068	**0.475**
Ngraph	0.298	−0.002	0.056	0.302	**0.421**
Rbook	0.212	0.019	0.014	−0.054	**0.414**

Rgraph	0.292	0.087	0.020	0.228	**0.393**
Rinstr	0.313	0.210	−0.058	0.161	**0.352**
Nmath	0.139	−0.037	0.023	0.205	**0.350**
Wreport	0.255	0.094	0.010	0.161	**0.339**
Present	0.159	0.226	0.035	0.034	**0.241**
ICTprogram	0.199	−0.045	0.037	0.023	**0.227**

Notes: The highest factor loading for each task item is indicated in bold. Exploratory factor analysis is performed based on 6,676 entrepreneur observations in PIAAC. The principal factor method and varimax orthogonal rotation are applied. The senate weights in PIAAC are applied. See Section 3.2 for additional details. For an explanation of each task item, see Table A1.

Table A4. Description and Summary Statistics of the Control Variables in the PIAAC Entrepreneur Regression (Column 1 Sample in Table 2)

Variable	Description	Total Mean	Total Std. dev.	Entrepreneurs Mean	Employees Mean
Female	Dummy: 1 if female, 0 if male	0.475	0.499	0.268	0.488
Age	Age	39.706	11.856	45.007	39.364
Exp	Years of work experience	17.822	11.782	22.935	17.492
Eduy	Years of education	13.401	2.773	13.359	13.404
LitScore	Logarithm of average literacy proficiency score based on PIAAC's direct skill assessment	5.618	0.163	5.600	5.619
NumScore	Logarithm of average numeric proficiency score based on PIAAC's direct skill assessment	5.614	0.180	5.616	5.614
Medu1	Dummy: 1 if mother's completed education is primary or lower secondary, 0 otherwise (reference category)	0.487	0.500	0.560	0.482
Medu2	Dummy: 1 if mother's completed education is upper secondary, 0 otherwise	0.346	0.476	0.304	0.349
Medu3	Dummy: 1 if mother's completed education is tertiary, 0 otherwise	0.167	0.373	0.137	0.169
Fedu1	Dummy: 1 if father's completed education is primary or lower secondary, 0 otherwise (reference category)	0.414	0.493	0.481	0.410
Fedu2	Dummy: 1 if father's completed education is upper secondary, 0 otherwise	0.377	0.485	0.323	0.380
Fedu3	Dummy: 1 if father's completed education is tertiary, 0 otherwise	0.209	0.407	0.196	0.210
Books	Number of books at home (midpoint number in the six-point range)	137.992	149.713	138.436	137.963
Htask_ICT	ICT-related factor score extracted from an exploratory factor analysis over the frequency of 25 tasks (activities) outside of work. The principal factor method and varimax orthogonal rotation are applied.	0.084	0.877	−0.043	0.092
Htask_Math	Math-related factor score similarly computed as *Htask_ICT*.	−0.149	0.707	−0.123	−0.150
Htask_Clerical	Clerical work-related factor score similarly computed as *Htask_ICT*.	−0.013	0.712	0.191	−0.026
LearnAttitude	Learning attitude index, which is the average of the respondent's standardized answers to each of six questions that measure the extent to which the following statements apply to the respondent: (i) *When I hear or read about new ideas, I try to relate them to real-life situations to which they might apply*, (ii) *I like learning new things*, (iii) *When I come across something new, I try to relate it to what I already know*, (iv) *I like to get to the bottom of difficult things*, (v) *I like to*	0.105	0.699	0.198	0.099

	figure out how different ideas fit together, and (vi) *If I don't understand something, I look for additional information to make it clearer*. Each answer is originally assessed on a 5-point scale (1= not at all to 5= to a very high extent). This index is essentially the same as the readiness-to-learn index constructed by OECD.				
Health	Health status (1 = *poor* to 5 = *excellent*)	3.551	0.977	3.502	3.555
HHsize	Number of people living in household	3.243	1.359	3.475	3.228
Livepartner	Dummy: 1 if living with a spouse or partner, 0 if not	0.674	0.469	0.830	0.664
Kids	Dummy: 1 if have children, 0 if not	0.657	0.475	0.823	0.646
Forborn	Dummy: 1 if born in a foreign country, 0 otherwise	0.109	0.311	0.097	0.109
Forlang	Dummy: 1 if PIAAC's test language is different from respondent's native language, 0 if the same	0.109	0.312	0.100	0.110
Trust	Answer to "There are only a few people you can trust completely" (1 = *strongly agree* to 5 = *strongly disagree*)	2.368	1.154	2.319	2.371
Altruism	Frequency of voluntary work in the last 12 months, approximated from a 5-point scale as follows: 1 = *never* (approximated as 0 per week), 2 = *less than once a month* (0.12), 3 = *less than once a week but at least once a month* (0.62), 4 = *at least once a week but not every day* (3), and 5 = *every day* (5).	0.869	1.488	1.078	0.855
Workhour	Hours of work per week	38.943	12.812	49.366	38.270
PrivateSector	Dummy: 1 if working in the private sector, 0 otherwise (reference category)	0.713	0.452	0.985	0.696
PublicSector	Dummy: 1 if working in the public sector, 0 otherwise	0.263	0.440	0.011	0.279
NPOSector	Dummy: 1 if working at a non-profit organization, 0 otherwise	0.024	0.152	0.004	0.025
Occ: Managers	Dummy: 1 if a manager, 0 otherwise	0.091	0.288	0.305	0.077
Occ: Professional	Dummy: 1 if a professional, 0 otherwise	0.211	0.408	0.142	0.215
Occ: Tech/ AssoPro	Dummy: 1 if a technician or associate professional, 0 otherwise	0.164	0.370	0.092	0.169
Occ: Clerks	Dummy: 1 if a clerk, 0 otherwise	0.111	0.314	0.033	0.116
Occ: Service/ Sales	Dummy: 1 if a service worker or shop and market sales worker, 0 otherwise	0.177	0.382	0.163	0.178
Occ: Skilled agri	Dummy: 1 if a skilled agricultural and fishery worker, 0 otherwise	0.013	0.112	0.094	0.007

Occ: Craft	Dummy: 1 if a craft and related-trades worker, 0 otherwise	0.101	0.301	0.131	0.099
Occ: Operator/ Assembler	Dummy: 1 if a plant and machine operator or assembler, 0 otherwise	0.070	0.254	0.028	0.072
Occ: Elementary	Dummy: 1 if an elementary occupation, 0 otherwise (reference category)	0.063	0.244	0.013	0.067
Industry 1	Agriculture, forestry, and fishing	0.020	0.140	0.108	0.014
Industry 2	Mining and quarrying	0.005	0.072	0.001	0.005
Industry 3	Manufacturing (reference category)	0.166	0.372	0.116	0.169
Industry 4	Electricity, gas, and water supply	0.015	0.123	0.007	0.016
Industry 5	Construction	0.064	0.245	0.135	0.060
Industry 6	Wholesale and retail trade/repair of motor vehicles and motorcycles	0.141	0.348	0.197	0.138
Industry 7	Transportation and storage	0.054	0.226	0.037	0.055
Industry 8	Accommodation and food service activities	0.048	0.214	0.090	0.045
Industry 9	Information and communication	0.038	0.190	0.028	0.038
Industry 10	Financial and insurance activities	0.035	0.184	0.019	0.036
Industry 11	Real estate activities/Administrative and support service activities	0.045	0.207	0.055	0.044
Industry 12	Professional, scientific, and technical activities	0.050	0.217	0.090	0.047
Industry 13	Public administration and defense/compulsory social security	0.068	0.252	0.003	0.072
Industry 14	Education	0.097	0.296	0.024	0.102
Industry 15	Human health and social work activities	0.114	0.318	0.042	0.119
Industry 16	Other services	0.040	0.195	0.050	0.039
BEL	Belgium (Flanders only)	0.042	0.200	0.048	0.042
CHL	Chile	0.032	0.176	0.048	0.031
CYP	Cyprus	0.028	0.166	0.024	0.029
CZE	Czech Republic	0.038	0.191	0.022	0.039
DEU	Germany	0.046	0.209	0.041	0.046
DNK	Denmark	0.050	0.219	0.036	0.051
ESP	Spain	0.033	0.178	0.035	0.033
EST	Estonia	0.044	0.206	0.051	0.044
FIN	Finland	0.045	0.208	0.040	0.045

FRA	France	0.031	0.172	0.025	0.031
GRC	Greece	0.023	0.148	0.056	0.020
IRL	Ireland	0.037	0.189	0.043	0.037
ISR	Israel	0.037	0.188	0.033	0.037
ITA	Italy	0.028	0.165	0.036	0.028
JPN	Japan	0.044	0.205	0.035	0.045
KOR	South Korea	0.038	0.191	0.078	0.035
LTU	Lithuania	0.038	0.192	0.018	0.040
NLD	Netherlands	0.049	0.216	0.046	0.049
NOR	Norway	0.046	0.210	0.022	0.048
NZL	New Zealand	0.045	0.207	0.065	0.043
POL	Poland	0.033	0.180	0.038	0.033
SGP	Singapore	0.044	0.205	0.049	0.044
SVK	Slovak Republic	0.033	0.178	0.024	0.033
SVN	Slovenia	0.035	0.183	0.022	0.035
SWE	Sweden	0.047	0.212	0.040	0.048
GBR	UK (England and Northern Ireland only, reference category)	0.034	0.181	0.022	0.035

Notes: Based on the PIAAC regression sample of Column (1) in Table 2. The number of observations is 75,883. Statistics are weighted by the senate weights in PIAAC. Most of the variable names and descriptions are taken from Table A1 of Asuyama (2022) ("Doing Boss-like Tasks and Worker Well-being: Job Enrichment Revisited," Yoko Asuyama, *Labour*, 2022, Volume 36, Issue 2, doi.org/10.1111/labr.12217, ©2022 Fondazione Giacomo Brodolini and John Wiley & Sons Ltd.).

Table A5. Description of Variables in Entrepreneurship Regressions on Learning Opportunities

Variable	Description	Unit	Source
Suskill	Dummy: 1 if the respondent answers yes to "*Do you have the knowledge, skill and experience required to start a new business?*," 0 if respondent answers no (see Section 4.2)	*icgaet*	GEM
EntrepY5	Dummy: 1 if the respondent is either a so-called "nascent entrepreneur" who is involved in setting up a business but not paying wages for the last 3 months or an "owner-manager of a new firm" that is less than 5 years old, 0 otherwise (see Section 4.2)	*icgaet*	GEM
TEA	Dummy: 1 if the respondent is involved in the GEM-defined "total early-stage entrepreneurial activity (TEA)," that is, either a nascent entrepreneur or an owner-manager of a new firm that is less than 3.5 years old, 0 otherwise (see Section 4.2)	*icgaet*	GEM
Emp5	Dummy: 1 if the "employment" (the number of people currently working for the respondent's business including owners and exclusive subcontractors) is 5 or more, 0 if it is 1–4 (see Section 5)	*icgaet*	GEM
Emp10	Dummy: 1 if the "employment" is 10 or more, 0 if it is 1–9 (see Section 5)	*icgaet*	GEM
Emp20	Dummy: 1 if the "employment" is 20 or more, 0 if it is 1–19 (see Section 5)	*icgaet*	GEM
Innov	Innovativeness of the respondent's business, which is computed as the average of the standardized answers (originally on a 3-point scale) to the following three questions: (i) "*Will all, some or none of your potential customers consider this product or service new and unfamiliar?*" (ii) "*Right now, are there many, few, or no other business offering the same products or services to your potential customers?*" (iii) "*How long have the technologies or procedures required for this product or service been available? Less than a year, between one and five years or longer than five years?*" (see Section 5)	*icgaet*	GEM
EntrLearn	Main index for employees' opportunity to learn entrepreneurial skills (see Sections 3.3 and 4.1)	*cgae*	PIAAC
Female	Dummy: 1 if female, 0 if male	*icgaet*	GEM
Age	Age	*icgaet*	GEM
Edu1 - Edu7	Dummies: 1 if the highest level of education completed is either pre-primary education (*Edu1*), primary or the first stage of basic education (*Edu2*), lower secondary or the second stage of basic education (*Edu3*), upper secondary education (*Edu4*), post-secondary non-tertiary education (*Edu5*), the first stage of tertiary education (*Edu6*), or the second stage of tertiary education (*Edu7*), 0 otherwise	*icgaet*	GEM
HHsize	Number of members making up the permanent household	*icgaet*	GEM
Knowent	Dummy: 1 if the respondent answers yes to "Do you know someone personally who started a business in the past 2 years?," 0 if the respondent answers no	*icgaet*	GEM
EntrOld	Population ratio (%) of old entrepreneurs (an owner-manager of a new firm founded over 10 years before year *t*)	*cgaet*	GEM

Fearfail	Dummy: 1 if the respondent answers yes to "Would fear of failure prevent you from starting a business?," 0 if the respondent answers no	*icgaet*	GEM
LearnAttitude	See Table A4	*cgae*	PIAAC
Htask_ICT	See Table A4	*cgae*	PIAAC
Htask_Math	See Table A4	*cgae*	PIAAC
Htask_Clerical	See Table A4	*cgae*	PIAAC
Fulltime	Ratio of full-time employees	*cgae*	PIAAC
Permanent	Ratio of employees with indefinite employment contracts	*cgae*	PIAAC
Estsize	Average number of employees at the workplace	*cgae*	PIAAC
Tenure	Workers' average years of tenure	*cgae*	PIAAC
Opport	Dummy: 1 if the respondent answers yes to "I*n the next six months, will there be good opportunities for starting a business in the area where you live?*," 0 if the respondent answers no	*icgaet*	GEM
Nbgoodc	Dummy: 1 if the respondent answers yes to "*In your country, most people consider starting a new business a desirable career choice?*," 0 if the respondent answers no	*icgaet*	GEM
Nbstatus	Dummy: 1 if the respondent answers yes to "*In your country, those successful at starting a new business have a high level of status and respect?*," 0 if the respondent answers no	*icgaet*	GEM
Nbmedia	Dummy: 1 if the respondent answers yes to "*In your country, you will often see stories in the public media and/or internet about successful new businesses?"* in 2014–2017, or "*In your country, you will often see stories in the public media about successful new businesses?*" in 2012–2013, 0 if the respondent answers no	*icgaet*	GEM
Trust	Workers' average trust, which is measured by the binary answer to *"Generally speaking, would you say that most people can be trusted* (1) *or that you need to be very careful in dealing with people* (0)*?*"	*cgae*	WVS-EVS
LOC	Workers' average locus of control, which is measured by how much freedom of choice and control the respondent feels that he or she has over the way his or her life turns out (1= no choice at all to 10 = a great deal of choice)	*cgae*	WVS-EVS
Creativity	Workers' average creativity focus, which is measured by how much the following person is like the respondent: "*It is important to this person to think up new ideas and be creative; to do things one's own way*" (1 = Not at all like the respondent to 6 = Very much like the respondent)	*cgae*	WVS-EVS
N-Ach	Workers' average need for achievement, which is measured by how much the following person is like the respondent: "*Being very successful is important to this person; to have people recognize one's achievements*" (1 = Not at all like the respondent to 6 = Very much like the respondent)	*cgae*	WVS-EVS
RiskPref	Workers' average risk preference, which is measured by how much the following person is like the respondent: "*Adventure and taking risks are important to this person; to have an exciting life*" (1 = Not at all like the respondent to 6 = Very much like the respondent)	*cgae*	WVS-EVS

Sources: PIAAC, GEM, WVS-EVS.

Table A6. Employees' Learning Opportunities, Skill Acquisition, and Occupational Choice: Results for the Control Variables

	LPM Model				**Probit model (AME)**			
Dep. Var.	*Suskill*		*EntrepY5*		*Suskill*		*EntrepY5*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
EntrLearn	2.452***	3.567***	0.901**	3.138***	2.825***	3.845***	0.762*	2.557***
	(0.676)	(1.115)	(0.432)	(0.741)	(0.660)	(1.111)	(0.392)	(0.693)
Female	-0.132***	-0.100***	-0.040***	-0.045***	-0.134***	-0.103***	-0.041***	-0.045***
	(0.006)	(0.012)	(0.004)	(0.010)	(0.006)	(0.012)	(0.003)	(0.008)
Age	0.012***	0.013***	0.005***	0.008***	0.002***	0.003***	0.000	0.000
	(0.002)	(0.002)	(0.001)	(0.002)	0.000	(0.001)	0.000	(0.001)
Age^2	-0.000***	-0.000***	-0.000***	-0.000***				
	0.000	0.000	0.000	0.000				
Edu1	-0.055***	-0.092***	0.007	0.015	-0.054***	-0.088***	0.009	0.021
	(0.018)	(0.024)	(0.010)	(0.016)	(0.017)	(0.023)	(0.012)	(0.018)
Edu2	-0.090***	-0.073***	0.004	0.000	-0.085***	-0.069***	0.003	0.001
	(0.010)	(0.013)	(0.007)	(0.008)	(0.010)	(0.012)	(0.008)	(0.009)
Edu3	-0.038***	-0.041***	-0.002	-0.008	-0.036***	-0.039***	-0.004	-0.009*
	(0.005)	(0.007)	(0.003)	(0.005)	(0.005)	(0.007)	(0.004)	(0.005)
Edu5	0.028***	0.023***	0.011***	0.005	0.027***	0.022***	0.011***	0.005
	(0.005)	(0.007)	(0.003)	(0.004)	(0.005)	(0.007)	(0.003)	(0.005)
Edu6	0.049***	0.033*	0.000	-0.006	0.052***	0.036**	0.003	0.004
	(0.011)	(0.017)	(0.007)	(0.012)	(0.011)	(0.017)	(0.007)	(0.012)
Edu7	0.036**	0.032	0.012	-0.023	0.037**	0.034	0.013	-0.008
	(0.015)	(0.025)	(0.012)	(0.017)	(0.015)	(0.024)	(0.010)	(0.015)
HHsize	0.003***	0.004**	0.001*	0.001	0.003***	0.004**	0.001**	0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Knowent	0.227***	0.214***	0.162***	0.149***	0.226***	0.212***	0.157***	0.143***
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
EntrOld	0.005***	0.006***	0.002***	0.004***	0.005***	0.006***	0.002***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	0.000	(0.001)
Fearfail	-0.135***	-0.125***	-0.082***	-0.085***	-0.134***	-0.124***	-0.079***	-0.081***
	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)
LearnAttitude	0.059*	0.057	0.025	0.011	0.069**	0.077**	0.013	-0.002
	(0.032)	(0.037)	(0.023)	(0.027)	(0.032)	(0.037)	(0.022)	(0.026)
Htask_ICT	-0.010	-0.007	0.009	-0.004	-0.021	-0.023	0.010	0.000
	(0.016)	(0.022)	(0.010)	(0.016)	(0.015)	(0.022)	(0.009)	(0.014)
Htask_Math	-0.059***	-0.075**	-0.002	-0.023	-0.067***	-0.086***	-0.006	-0.027
	(0.019)	(0.035)	(0.012)	(0.024)	(0.018)	(0.033)	(0.011)	(0.021)
Htask_Clerical	-0.045*	-0.029	0.002	-0.004	-0.048**	-0.029	0.005	-0.001
	(0.023)	(0.030)	(0.016)	(0.025)	(0.023)	(0.029)	(0.015)	(0.023)
Fulltime		0.051**		-0.033*		0.043*		-0.015
		(0.026)		(0.020)		(0.024)		(0.017)
Permanent		0.008		-0.072*		-0.030		-0.091**
		(0.059)		(0.040)		(0.055)		(0.038)
Estsize		-0.000***		0.000		-0.000**		0.000
		0.000		0.000		0.000		0.000

(*Table A6 continued*)

Dep. Var.	LPM Model				Probit model (AME)			
	Suskill		*EntrepY5*		*Suskill*		*EntrepY5*	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Tenure		0.000		0.000		0.001		0.000
		(0.002)		(0.001)		(0.002)		(0.001)
Opport		0.090***		0.062***		0.086***		0.058***
		(0.007)		(0.004)		(0.007)		(0.004)
Nbgoodc		0.001		0.001		0.002		0.003
		(0.005)		(0.003)		(0.005)		(0.003)
Nbstatus		-0.023***		-0.011***		-0.024***		-0.012***
		(0.005)		(0.004)		(0.005)		(0.003)
Nbmedia		0.003		0.004		0.003		0.003
		(0.005)		(0.003)		(0.005)		(0.003)
Trust		0.016		0.031		0.010		0.016
		(0.038)		(0.025)		(0.036)		(0.021)
LOC		0.007		0.013*		0.003		0.008
		(0.010)		(0.008)		(0.010)		(0.007)
Creativity		0.002		-0.006		0.007		-0.011
		(0.018)		(0.014)		(0.017)		(0.012)
N-Arch		0.050***		0.007		0.036**		-0.001
		(0.017)		(0.012)		(0.017)		(0.011)
RiskPref		0.023		0.022*		0.025*		0.019*
		(0.015)		(0.011)		(0.015)		(0.010)
Observations	251,544	122,846	237,074	113,096	251,544	122,846	237,074	113,096
# of clusters	442	224	442	224	442	224	442	224

Notes: The same as Table 4. Only results for Columns (2) and (5) in *Suskill* and *EntrepY5* regressions in Table 4 are reported. The results for year and country dummies are suppressed. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A7. Employees' Learning Opportunities, Skill Acquisition, and Occupational Choice: Using Alternative Learning Opportunity Indices

Dep. Var.	*Suskill*					*EntrepY5*					*TEA*				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
LPM model															
EntrLearnSQ	2.788***	2.058***	2.415***	1.976**	3.383***	1.191***	0.796**	1.144***	1.468***	2.619***	1.040***	0.702*	1.038**	1.388***	2.559***
	(0.572)	(0.647)	(0.703)	(0.785)	(0.984)	(0.376)	(0.395)	(0.431)	(0.537)	(0.660)	(0.362)	(0.377)	(0.406)	(0.511)	(0.648)
EntrLearnF	3.381***	2.734***	3.183***	2.545***	3.998***	1.442***	1.028**	1.509***	1.869***	3.324***	1.245***	0.887**	1.357***	1.748***	3.204***
	(0.612)	(0.680)	(0.745)	(0.832)	(1.128)	(0.400)	(0.433)	(0.479)	(0.588)	(0.730)	(0.391)	(0.419)	(0.458)	(0.567)	(0.727)
EntrLearnOB	3.165***	2.452***	2.882***	2.178***	3.567***	1.327***	0.901**	1.374***	1.660***	3.138***	1.140***	0.771*	1.234***	1.554***	3.056***
	(0.615)	(0.676)	(0.735)	(0.819)	(1.115)	(0.405)	(0.432)	(0.474)	(0.581)	(0.741)	(0.395)	(0.418)	(0.454)	(0.561)	(0.736)
EntrLearnL	2.960***	2.818***	3.180***	2.329***	3.696***	1.494***	1.054**	1.392***	1.690***	2.847***	1.313***	0.946**	1.277***	1.627***	2.878***
	(0.635)	(0.679)	(0.748)	(0.838)	(1.126)	(0.399)	(0.448)	(0.498)	(0.599)	(0.785)	(0.386)	(0.431)	(0.479)	(0.575)	(0.766)
EntrLearn40	2.938***	2.380***	2.787***	2.053**	3.608***	1.416***	0.946**	1.359***	1.661***	3.115***	1.216***	0.807**	1.225***	1.544***	3.043***
	(0.603)	(0.668)	(0.728)	(0.810)	(1.047)	(0.390)	(0.415)	(0.457)	(0.554)	(0.718)	(0.382)	(0.400)	(0.434)	(0.532)	(0.702)
Probit model (Average marginal effects [AME])															
EntrLearnSQ	3.062***	2.481***	2.846***	2.450***	3.763***	1.072***	0.685*	0.903**	1.170**	2.271***	0.910***	0.583*	0.793**	1.074**	2.132***
	(0.568)	(0.636)	(0.691)	(0.764)	(0.983)	(0.338)	(0.372)	(0.397)	(0.482)	(0.646)	(0.321)	(0.348)	(0.362)	(0.446)	(0.618)
EntrLearnF	3.513***	3.006***	3.439***	2.800***	4.077***	1.265***	0.863**	1.159***	1.435***	2.700***	1.072***	0.732**	1.013***	1.310***	2.515***
	(0.604)	(0.665)	(0.728)	(0.809)	(1.121)	(0.347)	(0.393)	(0.424)	(0.515)	(0.684)	(0.333)	(0.370)	(0.392)	(0.482)	(0.659)
EntrLearnOB	3.374***	2.825***	3.241***	2.555***	3.845***	1.176***	0.762*	1.068**	1.287**	2.557***	0.990***	0.638*	0.930**	1.168**	2.401***
	(0.608)	(0.660)	(0.718)	(0.796)	(1.111)	(0.354)	(0.392)	(0.417)	(0.509)	(0.693)	(0.339)	(0.368)	(0.386)	(0.476)	(0.665)
EntrLearnL	3.109***	3.155***	3.530***	2.750***	4.107***	1.195***	0.804*	1.123**	1.343**	2.343***	1.025***	0.713*	1.019**	1.285**	2.287***
	(0.620)	(0.657)	(0.724)	(0.817)	(1.123)	(0.356)	(0.415)	(0.437)	(0.534)	(0.743)	(0.339)	(0.392)	(0.411)	(0.502)	(0.711)
EntrLearn40	3.153***	2.768***	3.145***	2.434***	3.887***	1.298***	0.847**	1.074***	1.300***	2.612***	1.097***	0.712**	0.943**	1.178**	2.474***
	(0.596)	(0.652)	(0.707)	(0.782)	(1.032)	(0.341)	(0.385)	(0.408)	(0.494)	(0.688)	(0.327)	(0.361)	(0.376)	(0.460)	(0.653)
Controls	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5
Observations	*252,188*	251,544	224,306	152,228	122,846	*237,709*	237,074	210,895	140,520	113,096	*234,860*	234,229	208,389	138,660	111,526

Notes: For each row, the same regressions are performed as in Table 4, except that each alternative learning opportunity index is used instead of the baseline *EntrLearn*. When using *EntrLearn40*, the learning index (which is the average of standardized *LearnOJT*, *LearnCWboss*, and *Infonew* [see Table A1] and not used when predicting *EntrLearn40*) is also controlled for. For an explanation of the alternative learning opportunity indices, see Section 4.3 and Appendix B. When using *EntrLearnL* in Column (1), the number of observations is 253,159 (when Y = *Suskill*), 238,709 (when Y = *EntrepY5*) and 235,851 (when Y = *TEA*). ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A8. Employees' Learning Opportunities, Skill Acquisition, and Occupational Choice: *cgaet*-level LS and 2SLS Estimations

Dep. Var.	*Suskill*					*EntrepY5*					*TEA*				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
(a) LS															
EntrLearn	2.741***	2.284***	2.571***	2.063***	3.033***	1.148***	0.744*	1.153***	0.905*	1.915***	0.947**	0.588	1.001**	0.790*	1.794***
	(0.615)	(0.687)	(0.753)	(0.771)	(1.072)	(0.384)	(0.406)	(0.434)	(0.463)	(0.609)	(0.368)	(0.390)	(0.411)	(0.447)	(0.588)
Knowent	0.364***	0.362***	0.345***	0.329***	0.376***	0.275***	0.275***	0.243***	0.221***	0.236***	0.262***	0.263***	0.229***	0.205***	0.222***
	(0.032)	(0.031)	(0.033)	(0.035)	(0.044)	(0.021)	(0.021)	(0.023)	(0.023)	(0.029)	(0.020)	(0.020)	(0.021)	(0.022)	(0.027)
EntrOld		0.005***	0.005***	0.005***	0.006***		0.002***	0.003***	0.003***	0.004***		0.002***	0.002***	0.003***	0.003***
		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
Fearfail	-0.180***	-0.176***	-0.170***	-0.175***	-0.184***	-0.113***	-0.112***	-0.131***	-0.125***	-0.133***	-0.111***	-0.110***	-0.131***	-0.125***	-0.138***
	(0.028)	(0.026)	(0.030)	(0.031)	(0.039)	(0.021)	(0.021)	(0.022)	(0.021)	(0.025)	(0.020)	(0.020)	(0.021)	(0.020)	(0.023)
R-squared	0.810	0.820	0.808	0.815	0.825	0.623	0.629	0.594	0.605	0.637	0.616	0.622	0.580	0.590	0.626
Observations	1,954	1,950	1,679	1,537	1,012	1,946	1,942	1,670	1,514	1,010	1,943	1,939	1,667	1,510	1,008
(b) 2SLS															
EntrLearn	4.507**	3.168	4.300	2.543	4.830	4.344***	4.928***	6.355***	5.976***	6.935***	3.712***	4.028***	5.322***	4.959***	5.933***
	(1.854)	(2.741)	(2.700)	(2.900)	(3.127)	(1.058)	(1.527)	(1.725)	(1.837)	(1.906)	(0.974)	(1.402)	(1.546)	(1.653)	(1.714)
R-squared	0.808	0.819	0.805	0.814	0.825	0.604	0.602	0.545	0.563	0.605	0.602	0.603	0.543	0.559	0.602
Observations	1,914	1,910	1,642	1,503	1,008	1,910	1,906	1,637	1,486	1,006	1,908	1,904	1,635	1,482	1,004
Endogeneity test	1.556	0.221	0.702	0.097	0.353	11.866***	7.815***	11.193***	9.0723***	8.577***	10.397***	6.080**	9.236***	7.238***	6.741***
Overidentification test	0.437	0.500	0.528	0.311	0.090	1.484	1.508	0.132	0.330	0.218	1.684	1.862	0.211	0.442	0.331
First-stage (Y = *EntrLearn*)															
Boss	0.018***	0.013***	0.015***	0.014***	0.014***	0.018***	0.013***	0.014***	0.013***	0.014***	0.018***	0.013***	0.015***	0.013***	0.014***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
BossDistance	-0.008***	-0.006**	-0.006**	-0.005*	-0.008**	-0.008***	-0.007**	-0.007**	-0.006*	-0.008**	-0.008***	-0.007**	-0.007**	-0.006*	-0.008**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
1st stage F stat.	43.265	20.272	16.197	12.390	12.610	42.902	19.957	15.601	11.558	11.716	42.824	20.065	15.662	11.536	11.650
Controls	C1"	C2"	C3"	C4"	C5"	C1"	C2"	C3"	C4"	C5"	C1"	C2"	C3"	C4"	C5"

Notes: All of the dependent, independent, and instrumental variables used in the individual-level regressions (Tables 4 and 5) are collapsed at the *cgaet* level. The control variables are the same as in Tables 4 and 5, except that they are collapsed at the *cgaet* level. The ratio of individuals with a tertiary education and higher is controlled for instead of the education dummies (seven categories). All estimations are weighted by the senate weights in GEM summed over all target working individuals in each *cgaet* cell. Standard errors clustered by *cgae* environment are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A9. Employees' Learning Opportunities and Entrepreneur Performance: Sample Selection Model with Entrepreneur (*TEA* = 1) Definition

	Linear selection model					Probit selection model (coefficients)				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Y = Emp5										
EntrLearn	-1.274	-1.923	-2.200	-1.978	-4.902**	-4.161	-6.220	-7.545	-6.802	-17.442**
	(1.276)	(1.344)	(1.415)	(1.744)	(2.304)	(4.320)	(4.694)	(5.028)	(5.952)	(8.339)
Rho	-0.124	-0.120	-0.118	-0.049	-0.033	-0.190	-0.183	-0.182	-0.067	-0.052
	(0.048)	(0.049)	(0.051)	(0.085)	(0.102)	(0.091)	(0.091)	(0.100)	(0.133)	(0.149)
Test for *Rho* = 0	6.603**	6.000**	5.226**	0.327	0.104	4.184**	3.814*	3.151*	0.254	0.119
Obs.	225,340	224,734	201,115	133,329	107,431	225,340	224,734	201,115	133,329	107,431
Obs. selected	14,213	14,185	12,263	8,907	7,472	14,213	14,185	12,263	8,907	7,472
Y = Emp10										
EntrLearn	0.757	1.102	0.874	1.818	-0.378	3.854	5.612	3.598	7.018	-3.902
	(0.844)	(0.945)	(0.919)	(1.145)	(1.487)	(5.114)	(5.830)	(5.843)	(6.833)	(9.699)
Rho	-0.088	-0.085	-0.093	-0.063	-0.061	-0.167	-0.159	-0.182	-0.109	-0.129
	(0.042)	(0.043)	(0.044)	(0.061)	(0.071)	(0.114)	(0.114)	(0.127)	(0.157)	(0.185)
Test for *Rho* = 0	4.227**	3.772*	4.350**	1.066	0.738	2.081	1.863	1.969	0.472	0.478
Obs.	225,340	224,734	201,115	133,329	107,431	225,340	224,734	201,115	133,329	107,431
Obs. selected	14,213	14,185	12,263	8,907	7,472	14,213	14,185	12,263	8,907	7,472
Y = Emp20										
EntrLearn	0.374	1.066	1.333**	2.240**	-0.258	0.685	6.429	9.080	14.292	-11.580
	(0.608)	(0.673)	(0.675)	(0.952)	(1.100)	(6.202)	(7.094)	(7.395)	(9.417)	(13.726)
Rho	-0.057	-0.056	-0.072	-0.054	-0.047	-0.135	-0.126	-0.187	-0.128	-0.166
	(0.045)	(0.045)	(0.041)	(0.055)	(0.068)	(0.145)	(0.147)	(0.161)	(0.195)	(0.227)
Test for *Rho* = 0	1.603	1.528	3.033*	0.974	0.489	0.842	0.717	1.284	0.424	0.516
Obs.	225,340	224,734	201,115	133,329	107,431	225,340	224,734	201,115	133,329	107,431
Obs. selected	14,213	14,185	12,263	8,907	7,472	14,213	14,185	12,263	8,907	7,472
Y = Innov										
EntrLearn	-0.561	0.093	0.161	-1.608	-4.628					
	(1.903)	(2.081)	(2.033)	(2.339)	(3.002)					
Rho	-0.083	-0.085	-0.134	-0.069	0.022					
	(0.049)	(0.049)	(0.048)	(0.074)	(0.133)					
Test for *Rho* = 0	2.843*	3.036*	7.747***	0.869	0.028					
Obs.	233,058	232,430	206,864	137,768	110,842					
Obs. selected	21,931	21,881	18,012	13,346	10,883					
Controls	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5

Notes: The same as Table 7, except that selected observations are based on *TEA* = 1 sample. ***$p <$ 0.01, **$p < 0.05$, *$p < 0.1$.

Table A10. Employees' Learning Opportunities and Entrepreneur Performance: Sample Selection Model with Entrepreneur (*EntrepY5* = 1) Definition and Alternative Learning Opportunity Indices

	Linear selection model					Probit selection model (coefficients)				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Y = Emp5										
EntrLearnSQ	0.119	-0.679	-0.913	-0.578	-3.631*	0.405	-2.013	-2.987	-1.960	-12.690*
	(1.061)	(1.122)	(1.176)	(1.529)	(2.144)	(3.578)	(3.910)	(4.150)	(5.122)	(7.622)
Test for *Rho* = 0	6.532**	5.810**	4.958**	0.201	0.000	4.132**	3.696*	2.977*	0.135	0.000
EntrLearnF	-0.188	-1.074	-1.156	-0.927	-3.392	-0.416	-3.102	-3.797	-3.178	-11.777
	(1.180)	(1.245)	(1.295)	(1.672)	(2.249)	(3.969)	(4.345)	(4.586)	(5.655)	(7.970)
Test for *Rho* = 0	6.532**	5.797**	4.945**	0.200	0.000	4.134**	3.692*	2.971*	0.134	0.000
EntrLearnOB	-0.183	-1.082	-1.214	-0.949	-3.686	-0.444	-3.167	-3.930	-3.190	-12.696
	(1.193)	(1.257)	(1.308)	(1.678)	(2.241)	(4.002)	(4.367)	(4.606)	(5.662)	(7.929)
Test for *Rho* = 0	6.532**	5.800**	4.950**	0.202	0.000	4.134**	3.693*	2.974*	0.136	0.001
EntrLearnL	-1.256	-1.781	-1.663	-0.757	-2.745	-3.656	-5.198	-4.770	-1.826	-8.984
	(1.201)	(1.299)	(1.343)	(1.781)	(2.278)	(3.962)	(4.470)	(4.742)	(6.036)	(8.020)
Test for *Rho* = 0	6.687***	5.741**	4.908**	0.201	0.000	4.230**	3.668*	2.955*	0.135	0.000
EntrLearn40	-0.599	-1.750	-1.846	-1.578	-4.182*	-1.679	-5.270	-5.902	-5.202	-14.453*
	(1.161)	(1.238)	(1.285)	(1.652)	(2.148)	(3.906)	(4.299)	(4.514)	(5.558)	(7.584)
Test for *Rho* = 0	6.572**	5.799**	4.924**	0.201	0.000	4.159**	3.691*	2.941*	0.132	0.000
Obs.	*227,962*	227,352	203,451	135,067	108,915	*227,962*	227,352	203,451	135,067	108,915
Obs. selected	*16,835*	16,803	14,599	10,645	8,956	*16,835*	16,803	14,599	10,645	8,956
Y = Emp10										
EntrLearnSQ	1.956***	2.006**	1.906**	2.885***	0.349	10.223**	10.466**	9.365*	12.628**	1.136
	(0.713)	(0.820)	(0.799)	(1.061)	(1.294)	(4.327)	(4.987)	(4.844)	(5.791)	(8.215)
Test for *Rho* = 0	7.559***	6.993***	6.731***	1.095	0.326	3.4034*	3.150*	2.910*	0.450	0.180
EntrLearnF	1.704**	1.594*	1.584*	2.700**	0.704	9.260**	8.681	7.869	12.620*	4.492
	(0.776)	(0.889)	(0.862)	(1.127)	(1.332)	(4.599)	(5.383)	(5.370)	(6.442)	(8.806)
Test for *Rho* = 0	7.557***	6.968***	6.744***	1.107	0.327	3.398*	3.141*	2.921*	0.451	0.179
EntrLearnOB	1.743**	1.640*	1.562*	2.640**	0.659	9.397**	8.823	7.689	12.216*	3.569
	(0.778)	(0.897)	(0.874)	(1.143)	(1.351)	(4.603)	(5.403)	(5.425)	(6.507)	(8.760)
Test for *Rho* = 0	7.546***	6.958***	6.736***	1.112	0.328	3.398*	3.139*	2.925*	0.457	0.179
EntrLearnL	1.215	1.311	1.236	2.323*	0.275	6.104	6.519	6.221	11.069	0.968
	(0.807)	(0.939)	(0.927)	(1.263)	(1.473)	(4.524)	(5.539)	(5.628)	(7.119)	(9.114)
Test for *Rho* = 0	7.776***	7.004***	6.737***	1.145	0.322	3.499*	3.179*	2.952*	0.480	0.178
EntrLearn40	1.526**	1.209	1.140	2.127*	0.224	8.161*	6.382	5.423	9.108	0.045
	(0.773)	(0.889)	(0.884)	(1.175)	(1.349)	(4.575)	(5.357)	(5.448)	(6.587)	(8.554)
Test for *Rho* = 0	7.608***	6.960***	6.711***	1.104	0.320	3.405*	3.119*	2.880*	0.444	0.173
Obs.	*227,962*	227,352	203,451	135,067	108,915	*227,962*	227,352	203,451	135,067	108,915
Obs. selected	*16,835*	16,803	14,599	10,645	8,956	*16,835*	16,803	14,599	10,645	8,956

(*Table A10 continued*)

	Linear selection model					Probit selection model (coefficients)				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Y = Emp20										
EntrLearnSQ	0.728	1.245**	1.504**	2.399***	0.404	4.739	8.666	11.352*	16.068**	-2.735
	(0.479)	(0.577)	(0.593)	(0.828)	(0.950)	(5.046)	(5.973)	(6.340)	(7.972)	(12.278)
Test for *Rho* = 0	2.739*	2.621	3.672*	2.399	1.206	1.192	1.083	1.432	0.935	0.664
EntrLearnF	0.435	0.845	1.213**	2.125**	0.103	2.184	5.297	8.873	15.030*	-5.510
	(0.519)	(0.610)	(0.615)	(0.847)	(0.934)	(5.413)	(6.472)	(6.890)	(8.678)	(12.153)
Test for *Rho* = 0	2.715*	2.583	3.648*	2.393	1.178	1.178	1.066	1.416	0.910	0.654
EntrLearnOB	0.560	1.026*	1.377**	2.331***	0.298	3.222	6.847	10.329	16.714*	-4.056
	(0.523)	(0.619)	(0.624)	(0.862)	(0.934)	(5.479)	(6.567)	(6.978)	(8.804)	(12.021)
Test for *Rho* = 0	2.720*	2.591	3.656*	2.410	1.190	1.179	1.066	1.416	0.915	0.659
EntrLearnL	0.146	0.636	0.799	1.879**	-0.335	-2.846	0.629	2.815	10.719	-16.479
	(0.590)	(0.680)	(0.702)	(0.954)	(1.050)	(5.959)	(7.048)	(7.551)	(9.587)	(12.152)
Test for *Rho* = 0	3.166*	2.601	3.642*	2.478	1.182	1.373	1.074	1.437	0.953	0.647
EntrLearn40	0.541	0.990	1.379**	2.436***	0.416	3.006	6.505	10.584	17.452**	-3.074
	(0.520)	(0.619)	(0.630)	(0.860)	(0.914)	(5.398)	(6.586)	(7.037)	(8.778)	(12.127)
Test for *Rho* = 0	2.743*	2.611	3.675*	2.423	1.189	1.194	1.073	1.425	0.920	0.656
Obs.	*227,962*	227,352	203,451	135,067	108,915	*227,962*	227,352	203,451	135,067	108,915
Obs. selected	*16,835*	16,803	14,599	10,645	8,956	*16,835*	16,803	14,599	10,645	8,956
Y = Innov										
EntrLearnSQ	-0.756	-0.228	-0.001	-1.280	-4.169					
	(1.582)	(1.687)	(1.628)	(1.942)	(2.689)					
Test for *Rho* = 0	3.239*	3.434*	8.874***	0.481	0.164					
EntrLearnF	-0.628	-0.064	-0.096	-1.440	-3.694					
	(1.714)	(1.855)	(1.811)	(2.153)	(2.834)					
Test for *Rho* = 0	3.254*	3.452*	8.883***	0.478	0.166					
EntrLearnOB	-0.208	0.473	0.433	-1.132	-3.284					
	(1.748)	(1.898)	(1.851)	(2.174)	(2.874)					
Test for *Rho* = 0	3.258*	3.460*	8.899***	0.480	0.163					
EntrLearnL	0.396	1.616	0.696	-1.361	-2.424					
	(1.766)	(2.061)	(2.076)	(2.489)	(3.057)					
Test for *Rho* = 0	3.380*	3.492*	8.883***	0.468	0.167					
EntrLearn40	-0.249	0.230	0.203	-1.164	-3.529					
	(1.701)	(1.866)	(1.829)	(2.153)	(2.792)					
Test for *Rho* = 0	3.271*	3.453*	8.876***	0.480	0.162					
Obs.	*235,584*	234,953	209,107	139,445	112,272					
Obs. selected	*24,457*	24,404	20,255	15,023	12,313					
Controls	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5

Notes: For each row, the same regressions are performed as in Table 7, except that each alternative learning opportunity index is used instead of *EntrLearn*. When using *EntrLearn40*, the learning index (which is the average of standardized *LearnOJT*, *LearnCWboss*, and *Infonew* [see Table A1] and not used when predicting *EntrLearn40*) is also controlled for. For an explanation of alternative learning opportunity indices, see Section 4.3 and Appendix B. When using E*ntrLearnL* in Column (1), the number of observations and those selected are 228,902 and 16,909 (when *Y* = *Emp5*, *Emp10*, *or Emp20*) and 236,573 and 24,580, respectively (when Y = *Innov*). ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A11. Employees' Learning Opportunities and Entrepreneurship Performance: Extended Regression Models for Entrepreneur (*EntrepY5* = 1) Sample

Dep. Var.	***Emp5***					***Emp10***				
Model	Extended probit regression (coefficients)					Extended probit regression (coefficients)				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Outcome equation (*Y* = Dep. Var.)										
EntrLearn	5.802	4.768	2.301	-0.342	0.747	9.945	15.182	2.549	-11.049	-1.196
	(10.040)	(15.578)	(18.524)	(23.931)	(24.061)	(10.474)	(17.401)	(21.901)	(27.691)	(26.650)
Selection equation (*Y* = *EntrepY5*)										
Fearfail	-0.476***	-0.476***	-0.483***	-0.462***	-0.489***	-0.476***	-0.476***	-0.483***	-0.462***	-0.489***
	(0.014)	(0.014)	(0.015)	(0.017)	(0.020)	(0.014)	(0.014)	(0.015)	(0.017)	(0.020)
Endogenous variable equation (*Y* = *EntrLearn*)										
Boss	0.018***	0.014***	0.015***	0.013***	0.015***	0.018***	0.014***	0.015***	0.013***	0.015***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
BossDistance	-0.008***	-0.006***	-0.006**	-0.005**	-0.006**	-0.008***	-0.006***	-0.006**	-0.005**	-0.006**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Test for selection	-0.163**	-0.155*	-0.151*	-0.041	0.001	-0.228**	-0.221**	-0.245**	-0.158	-0.100
Test for endogeneity	-0.027	-0.028	-0.023	-0.015	-0.038	-0.007	-0.024	0.014	0.070	0.013
Obs.	226,670	226,081	202,262	134,212	108,786	226,670	226,081	202,262	134,212	108,786
Obs. selected	16,689	16,658	14,459	10,521	8,940	16,689	16,658	14,459	10,521	8,940
Controls	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5

Dep. Var.	***Emp20***					***Innov***				
Model	Extended probit regression (coefficients)					Extended linear regression				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Outcome equation (*Y* = Dep. Var.)										
EntrLearn	-10.819	-3.121	-23.506	-34.298	-44.232	1.670	6.984	3.438	5.858	8.390
	(14.332)	(21.284)	(24.816)	(32.318)	(34.993)	(3.912)	(6.253)	(7.560)	(9.661)	(10.011)
Selection equation (*Y* = *EntrepY5*)										
Fearfail	-0.476***	-0.476***	-0.483***	-0.462***	-0.489***	-0.467***	-0.467***	-0.481***	-0.451***	-0.469***
	(0.014)	(0.014)	(0.015)	(0.017)	(0.020)	(0.013)	(0.012)	(0.014)	(0.015)	(0.018)
Endogenous variable equation (*Y* = *EntrLearn*)										
Boss	0.018***	0.014***	0.015***	0.013***	0.015***	0.018***	0.014***	0.015***	0.013***	0.015***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
BossDistance	-0.008***	-0.006***	-0.006**	-0.005**	-0.006**	-0.008***	-0.006***	-0.006**	-0.005**	-0.006**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Test for selection	-0.197	-0.192	-0.251*	-0.260	-0.191	-0.076*	-0.077*	-0.123***	-0.039	0.072
Test for endogeneity	0.051	0.035	0.114	0.162*	0.116	-0.010	-0.031	-0.013	-0.032	-0.050
Obs.	226,670	226,081	202,262	134,212	108,786	234,223	233,617	207,857	138,550	112,133
Obs. selected	16,689	16,658	14,459	10,521	8,940	24,242	24,194	20,054	14,859	12,287
Controls	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5

Notes: The extended regression model (Stata Press 2021) that deals with both selection and endogenous covariates is estimated by applying the Stata commands "eprobit" and "eregress." For more information on the dependent variables, the exclusion restriction variable (*Fearfail*), and the IVs for *EntrLearn* (*Boss* and *BossDistance*), see Tables 5 and 7. *Knowent* is controlled in all columns and *EntrOld* is controlled for in Columns (2)–(5). For the other control variables, see Table 4. The "test for selection" reports the estimated correlation of the error terms of the selection and outcome equations. The "test for endogeneity" reports the estimated correlation of the error terms of the *EntrLearn* and outcome equations. Statistically significant correlations respectively imply the existence of selection bias or endogeneity bias of *EntrLearn*. All estimations are weighted by the senate weights in GEM. Standard errors clustered by *cgae* environment are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A12. Regression-Based Variance Decomposition of Employee-Level Learning Opportunities for Entrepreneur Tasks ($EntrLearn_i$)

(Units: %)

	Boss = $Boss_i$	Boss = {# of subordinates, $BossDistance_i$}
	(1)	(2)
Residual	77.6	77.5
Boss	2.6	2.6
Skill	1.8	1.8
Gender	0.2	0.2
Age	0.3	0.2
Exp	0.7	0.7
Foreign	0.2	0.2
PPsector	1.3	1.3
Occ	9.4	9.5
Industry	3.8	3.8
Country	2.3	2.3

Notes: The regression-based variance decomposition method (Fields 2003) is applied to PIAAC employee data. The figures in this table present the percentage of variance of the dependent variable ($EntrLearn_i$) explained by each element. *Boss* includes a boss dummy ($Boss_i$ used in Section 4.4) in Column (1) or the number of subordinates [six categories] and $BossDistance_i$ (used in Section 4.4) in Column (2). *Skill* includes the years of education and literacy and numeric proficiency scores. *Gender* is a female dummy. *Age* includes age and its square. *Experience* includes years of work experience and its square. *Foreign* includes a foreign-born dummy and a foreign native language dummy. *PPsector* includes dummies for the public sector and for the non-profit organization sector (with the private sector as the reference category). *Occupation* includes occupation dummies (10 categories). Industry includes industry dummies (16 categories). Country includes country dummies (19 countries).

Appendix B: Construction of Alternative Learning Opportunity Indices

This appendix explains the construction of the alternative learning opportunity indices ($EntrLearnOB_{cgae}$, $EntrLearnL_{cgae}$, and $EntrLearn40_{cgae}$) mentioned in Section 4.3. For all indices, explanatory factor analysis (EFA) is performed based on all entrepreneur observations in PIAAC for 31 countries. PIAAC's senate weights are applied. Factor scores are calculated using the regression scoring method, as is the case when calculating the main $EntrLearn_{cgae}$.

$\boldsymbol{EntrLearnOB_{cgae}}$: This index is the employees' *cgae*-cell average of $EntrLearnOB_i$, computed using Equation (4) with an alternative vector of task factor scores, $\mathbf{TaskF}_i$. This $\mathbf{TaskF}_i$ is extracted from the same EFA in Section 3.2, except that a quartimin oblique rotation, which allows correlations between the factors, is used instead of a varimax orthogonal rotation. The factor loadings of 47 task items after quartimin oblique rotation are reported in Table B1(a).

$\boldsymbol{EntrLearnL_{cgae}}$: This index is the employees' *cgae*-cell average of $EntrLearnL_i$, computed using Equation (4) with an alternative vector of task factor scores, $\mathbf{TaskF}_i$. This $\mathbf{TaskF}_i$ is extracted from the same EFA in Section 3.2, except that low-loading and multi-loading task items are dropped when constructing the factor score. Following Tabachnick and Fidell (2007: 649), task items with the highest loading less than 0.32 are considered to be low-loading items. Task items that load more than 0.32 for multiple factors are considered multi-loading items.[1] Consequently, 13 task items are dropped. Table B1(b) reports the factor loadings of the remaining 34 task items.

$\boldsymbol{EntrLearn40_{cgae}}$: This index is the employees' *cgae*-cell average of $EntrLearn40_i$, which is computed as:

---

[1] The EFA is performed by taking the following steps, which aim to minimize the number of task items dropped: In Step 1, starting with five factors, EFA is performed repeatedly by dropping task items with loadings less than 0.3. In Step 2, if there are no task items with a loading less than 0.3, items with loadings less than 0.32 are dropped. Steps 1 and 2 are repeated until no low-loading items (i.e., items with a loading of less than 0.32) remain. In Step 3, multi-loading items that load more than 0.32 for multiple factors are dopped.

$$EntrLearn40_i = \Phi(\widehat{\alpha_2} + \mathbf{TaskF_i}\widehat{\boldsymbol{\beta_2}} + \widehat{\gamma_2} TaskDiversity_i + \widehat{\varphi_2} Auto_i + \widehat{\theta_2}\overline{Learn_i} + \overline{\mathbf{X}}\widehat{\boldsymbol{\delta_2}}), \qquad \text{(B1)}$$

where $\widehat{\alpha_2}$, $\widehat{\boldsymbol{\beta_2}}$, $\widehat{\varphi_2}$, $\widehat{\gamma_2}$, and $\widehat{\boldsymbol{\delta_2}}$ are estimated from the following probit model:

$$\mathrm{P}(Y = 1|\mathbf{Z}) = \Phi(\alpha_2 + \mathbf{TaskF}\boldsymbol{\beta_2} + \gamma_2 TaskDiversity + \varphi_2 Auto + \theta_2 Learn + \mathbf{X}\boldsymbol{\delta_2}). \text{(B2)}$$

TaskF is a vector of task factor scores extracted from the same EFA in Section 3.2, except that EFA is performed based on 40 task items instead of 47 items. Four autonomy-related task items (*AutoWay*, *AutoOrder*, *AutoSpeed*, *AutoHour* in Table A1) and three learning-related task items (*LearnOJT*, *LearnCWboss*, *Infonew* in Table A1) are not included in the EFA because they are related to the entirety of the respondent's work and therefore they are different from more specific tasks. The factor loadings of the 40 tasks are presented in Table B1(c). *Auto* and *Learn*, respectively, is the average of four standardized autonomy-related items (*AutoWay*, *AutoOrder*, *AutoSpeed*, *AutoHour*) and three standardized learning-related items (*LearnOJT*, *LearnCWboss*, *Infonew*). As Equation (B1) shows, individual-level $Learn_i$ is not used when predicting $EntrLearn40_i$. Instead, sample average $Learn_i$ ($\overline{Learn_i}$) is used. This is because even though $\widehat{\theta_2}$ turns out to be negative, assuming that more learning impedes entrepreneurship seems to be an odd conclusion. The employees' *cgae*-cell average of $Learn_i$ is also controlled for when estimating the effect of $EntrLearn40_{cgae}$ on entrepreneurial activities in Tables A7 and A10.

Additional References

Tabachnick, Barbara G., and Linda S. Fidell. 2007. *Using Multivariate Statistics, Fifth Edition*. Boston: Pearson Education.

Table B1. Factor Loadings of the Task Items Used When Constructing Alternative Learning Opportunity Indices

(a) Factor Loadings of 47 Task Items in $EntrLearnOB_{cgae}$: Quartimin Oblique Rotation is Applied

Factor	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor name	*ICT*	*Mgmt*	*Finance*	*Clerical*	*Autonomy*
Task item					
ICTmail	**0.938**	0.033	−0.055	−0.086	−0.018
ICTinternet	**0.779**	0.026	−0.025	0.068	−0.007
Rmail	**0.751**	0.091	−0.015	0.010	−0.001
Wmail	**0.742**	0.048	0.042	0.032	0.024
ICTword	**0.709**	−0.062	0.062	0.112	0.039
ICTec	**0.665**	0.007	0.120	−0.060	0.012
ICTexcel	**0.575**	−0.106	0.196	0.108	0.027
Manuwork	**−0.462**	0.264	0.035	−0.041	−0.085
Rnews	**0.373**	0.033	−0.043	0.274	0.026
ICTchat	**0.262**	0.001	−0.013	0.143	0.025
Planother	0.017	**0.551**	0.121	−0.059	0.036
Infoshare	0.062	**0.534**	−0.028	0.041	0.019
Planown	0.077	**0.526**	0.019	−0.148	0.071
Teach	−0.002	**0.504**	−0.030	0.187	0.013
Persuade	0.111	**0.499**	0.164	0.035	0.006
Advise	0.180	**0.483**	0.131	0.008	0.018
Cooperate	−0.158	**0.460**	0.001	0.043	−0.016
Mnghour	0.067	**0.455**	0.014	−0.146	0.122
Negotiate	0.076	**0.455**	0.261	0.042	−0.021
PSeasy	0.039	**0.404**	0.109	0.130	−0.003
LearnOJT	−0.034	**0.384**	−0.137	0.281	0.011
LearnCWboss	0.033	**0.355**	−0.121	0.295	0.021
Infonew	0.076	**0.327**	0.003	0.279	0.037
PSdiff	0.057	**0.311**	0.034	0.291	−0.002
Fingerwork	−0.202	**0.299**	0.049	0.042	−0.106
Nprice	−0.069	0.103	**0.693**	−0.033	−0.016
Nfraction	0.095	0.028	**0.613**	0.118	0.009
Ncalculator	0.128	0.029	**0.594**	−0.026	0.016
Nalgebra	−0.004	−0.051	**0.421**	0.220	0.058
Rfinstat	0.235	0.102	**0.383**	−0.059	0.010
Sales	−0.051	0.297	**0.352**	−0.157	0.011
Rmanual	0.074	0.142	0.006	**0.493**	−0.043
Rbook	0.079	−0.000	−0.098	**0.446**	0.016
Ngraph	0.065	−0.086	0.295	**0.439**	0.054
Rjournal	0.298	0.090	−0.064	**0.408**	0.015
Rgraph	0.083	0.023	0.200	**0.403**	0.013
Nmath	−0.045	−0.093	0.210	**0.377**	0.032
Rinstr	0.135	0.168	0.106	**0.348**	−0.075

Wreport	0.084	0.047	0.128	**0.347**	0.002
Wform	0.115	0.042	0.247	**0.266**	−0.022
ICTprogram	0.124	−0.076	0.005	**0.242**	0.032
Present	0.032	0.220	−0.021	**0.238**	0.027
Wnews	0.148	0.005	−0.042	**0.183**	0.024
AutoWay	−0.009	0.031	−0.010	0.021	**0.818**
AutoOrder	0.015	0.020	0.006	−0.029	**0.810**
AutoSpeed	−0.043	0.001	0.009	−0.014	**0.805**
AutoHour	0.027	−0.051	−0.026	0.018	**0.703**

(b) Factor Loadings of 34 Task Items in $EntrLearnL_{cgae}$:
Low−Loading and Multi−Loading Items Dropped

Factor	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor name	*ICT*	*Finance*	*Clerical*	*Mgmt*	*Autonomy*
Task item					
ICTmail	**0.913**	−0.058	−0.035	0.059	−0.015
ICTword	**0.773**	0.023	−0.002	−0.012	0.021
ICTinternet	**0.755**	−0.008	0.109	0.005	−0.004
Wmail	**0.738**	0.043	0.043	0.045	0.026
Rmail	**0.704**	0.009	0.114	0.049	0.012
ICTexcel	**0.658**	0.140	−0.052	−0.029	0.003
ICTec	**0.650**	0.119	−0.048	0.036	0.012
Manuwork	**−0.440**	0.044	0.011	0.213	−0.081
Nprice	−0.092	**0.750**	−0.019	0.025	−0.011
Ncalculator	0.090	**0.654**	0.004	−0.044	0.024
Nfraction	0.132	**0.630**	0.046	−0.022	0.001
Sales	−0.147	**0.443**	−0.002	0.202	0.024
Rfinstat	0.200	**0.420**	−0.002	0.051	0.020
Nalgebra	0.112	**0.354**	0.046	−0.029	0.031
Rmanual	0.073	0.068	**0.534**	−0.030	−0.038
Rjournal	0.273	−0.003	**0.479**	−0.066	0.027
LearnCWboss	−0.044	−0.010	**0.464**	0.174	0.029
Infonew	−0.014	0.128	**0.462**	0.121	0.053
LearnOJT	−0.101	−0.029	**0.460**	0.185	0.023
Rbook	0.086	−0.056	**0.446**	−0.128	0.021
Rnews	0.299	0.031	**0.400**	−0.113	0.049
Rinstr	0.152	0.139	**0.368**	0.036	−0.070
Planown	0.121	−0.034	−0.109	**0.618**	0.029
Planother	0.052	0.098	−0.013	**0.581**	0.000
Mnghour	0.100	−0.036	−0.104	**0.536**	0.087
Infoshare	0.039	−0.001	0.170	**0.505**	−0.003
Cooperate	−0.184	0.036	0.155	**0.423**	−0.032
Teach	0.007	0.015	0.257	**0.420**	−0.010
Advise	0.135	0.185	0.150	**0.393**	0.012
Persuade	0.076	0.221	0.176	**0.350**	0.008
AutoWay	−0.012	0.003	0.035	0.010	**0.820**

AutoOrder	0.007	0.016	−0.010	0.013	**0.811**
AutoSpeed	−0.039	0.009	−0.016	0.003	**0.805**
AutoHour	0.045	−0.042	−0.012	−0.034	**0.700**

(c) Factor Loadings of 40 Task Items in $EntrLearn40_{cgae}$:
Four Autonomy-Related and Three Learning-Related Task Items Not Covered

Factor	Factor 1	Factor 2	Factor 3	Factor 4
Factor name	*ICT*	*Mgmt*	*Finance*	*Clerical*
Task item				
ICTmail	**0.879**	0.111	0.060	−0.002
ICTinternet	**0.788**	0.096	0.083	0.121
Wmail	**0.766**	0.139	0.129	0.110
Rmail	**0.752**	0.166	0.083	0.090
ICTword	**0.752**	0.041	0.124	0.169
ICTec	**0.670**	0.093	0.185	0.014
ICTexcel	**0.646**	0.003	0.230	0.160
Rnews	**0.459**	0.086	0.014	0.293
Rjournal	**0.437**	0.134	0.000	0.410
Manuwork	**−0.435**	0.210	0.009	−0.028
ICTchat	**0.315**	0.030	0.030	0.140
Wnews	**0.203**	0.032	−0.021	0.184
Planother	0.143	**0.580**	0.170	0.041
Planown	0.152	**0.537**	0.073	−0.063
Infoshare	0.177	**0.521**	0.067	0.092
Persuade	0.265	**0.518**	0.251	0.105
Advise	0.312	**0.517**	0.213	0.093
Teach	0.155	**0.499**	0.053	0.229
Negotiate	0.247	**0.478**	0.340	0.104
Mnghour	0.141	**0.469**	0.059	−0.072
Cooperate	−0.054	**0.423**	0.065	0.067
PSeasy	0.194	**0.400**	0.197	0.162
PSdiff	0.232	**0.304**	0.128	0.285
Fingerwork	−0.145	**0.277**	0.060	0.080
Nprice	0.131	0.186	**0.673**	0.026
Nfraction	0.315	0.130	**0.607**	0.167
Ncalculator	0.294	0.120	**0.592**	0.030
Rfinstat	0.338	0.177	**0.402**	0.010
Nalgebra	0.198	0.036	**0.393**	0.236
Sales	0.056	0.309	**0.385**	−0.110
Rmanual	0.262	0.160	0.067	**0.473**
Ngraph	0.298	0.011	0.273	**0.442**
Rgraph	0.282	0.107	0.195	**0.428**
Rbook	0.208	0.026	−0.072	**0.423**
Wreport	0.242	0.122	0.119	**0.385**
Rinstr	0.293	0.204	0.157	**0.370**

Nmath	0.139	−0.030	0.183	**0.361**
Wform	0.269	0.131	0.233	**0.322**
Present	0.157	0.239	0.017	**0.254**
ICTprogram	0.204	−0.044	0.017	**0.226**

Notes: The highest factor loading for each task item is indicated in bold. See further explanation in Appendix B.