Learning Entrepreneurship from Other Entrepreneurs?

Luigi Guiso, *Einaudi Institute for Economics and Finance (EIEF)*

Luigi Pistaferri, *Stanford University*

Fabiano Schivardi, *Luiss University and EIEF*

We document that individuals who grow up in high firm density areas are more likely to become entrepreneurs, given firm density in their current location, and to run businesses in the sector with the highest density when young. Firm density at an entrepreneur's young age drives current firm profitability and is more important than current density for business performance. Results hold in a sample of movers, which allows addressing endogeneity concerns. These results are consistent with entrepreneurial skills being partly learnable through social contacts. Accordingly, entrepreneurs who grow up in high firm density areas adopt better managerial practices.

I. Introduction

Who becomes an entrepreneur? The answer economists give to this question is that individuals choose their occupation by comparing the costs and benefits of alternative occupations. In the classical Lucas (1978)/Rosen (1982)

We thank Michael Fritsch, Ed Glaeser, Vernon Henderson, Mirjam Van Praag, Antoinette Schoar, and Matt Turner as well as participants in the National Bureau of Economic Research (NBER) Summer Institute, the Fifth HEC Workshop on Entrepreneurship–Entrepreneurial Finance, the Second Center for Economic and Policy Research (CEPR) European Workshop on Entrepreneurship Economics, the Third Empirical Management Conference, and the Fifth Institute of Labor Economics (IZA) Workshop on Entrepreneurship Research for useful comments and

[*Journal of Labor Economics*, 2021, vol. 39, no. 1]
 0734-306X/2021/3901-0005$10.00
Submitted January 14, 2019; Accepted January 31, 2020; Electronically published October 16, 2020

model of occupational choice, individuals with greater managerial skills—defined as the ability to extract more output from a given combination of capital and labor—will sort into entrepreneurship because the return from managing a firm exceeds the wage they can earn as employees. Entrepreneurial skills can be interpreted more broadly to include, for instance, the ability to manage (and stand) risk and the capacity to identify and assess the economic potential of a new product or process. While it remains unclear how people obtain these skills, it is important to understand whether such skills are innate characteristics or are instead acquired through learning—and if so, how.

Distinguishing between these two sources of entrepreneurial skills (i.e., innate or learned) has important policy implications. If entrepreneurial ability is innate, then its distribution should not differ substantially across populations, and much of the observed differences in entrepreneurship across countries or regions within countries should be traced back to factors that facilitate or discourage people with entrepreneurial abilities to set up a firm—such as the availability of capital or institutional frictions. Fostering entrepreneurship thus requires removing these obstacles. If instead managerial abilities can be acquired through learning, differences in entrepreneurship can partly reflect differences in learning opportunities across countries or regions and the constraints to entrepreneurship are induced by learning frictions. Fostering entrepreneurship requires improving the learning process.

In this paper we investigate whether selection into entrepreneurship and entrepreneurial success are affected by learning opportunities. While individuals can learn how to become an entrepreneur and how to be a successful one in a variety of ways (e.g., from parents, friends, and schools) and at different stages of their life cycle, we look at one specific channel: learning from one's environment during the formative years (adolescence). Arguably, for a young individual growing up in Silicon Valley it should be easier to learn how to set up a firm (an entry cost channel) and how to run it (an entrepreneurial/managerial skill channel) than it would be for those growing up elsewhere. This is because the high concentration of entrepreneurial activities in the area provides many direct or indirect learning opportunities—both before and after entering the labor market, implying that the learning process can potentially occur both from and within firms (as a worker, apprentice, etc.) before a personal business is formally set up. And since firm density in the Silicon Valley region is highest in the information and communication technology (ICT) sector, a young individual should be more

suggestions. Financial support from the Regione Sardegna, Legge 7/2007 (grant CRP 26133), is gratefully acknowledged. Contact the corresponding author, Luigi Pistaferri, at pista@stanford.edu. Information concerning access to the data used in this paper is available as supplemental material online.

likely to start a business in this sector either because she learned more specific skills or because she learned how to enter it at a lower cost. We study whether these intuitive predictions receive empirical support. In particular, we test whether firm density in the location where individuals grow up affects the choice of becoming an entrepreneur, the sector of activity chosen, and their subsequent performance as entrepreneurs.

We analyze the effects of learning opportunities on entrepreneurial choices and outcomes using a simple extension of the Lucas (1978) occupational choice model. We let learning opportunities operate along two dimensions. First, individuals can learn skills that are useful to run a business (the "skill improvement" channel); second, they can learn how to set up a business efficiently (the "entry cost reduction" channel). We show that these two channels have the same implications in terms of occupational choices—individuals are more likely to become entrepreneurs when there are more learning opportunities, independently of the channel. However, they affect the performance of entrepreneurs in opposite directions: the skill improvement channel implies that on average entrepreneurs are more capable the higher learning opportunities are, while the entry cost channel reduces average entrepreneurial performance.

We study the effects of learning opportunities using a variety of data sets. The first is a sample of Italian entrepreneurs actively managing a small or medium-sized incorporated firm (the Associazione Nazionale delle Imprese Assicurative [ANIA] sample). Besides a rich set of demographic variables, this database contains detailed information on the sector of the firm and the entrepreneurs' place of birth, current location, and location at age 18 (which we term the "learning age"). We match these entrepreneurs with their firms' balance sheet data and thus obtain measures of firm total factor productivity (TFP) and sales per worker. This allows us to test two of the implications of the learning model: first, conditional on becoming an entrepreneur, one should be more likely to start a firm in a sector where density at learning age was particularly high; second, firms' productivity should be increasing with the firm density of the location in which the entrepreneur lived at learning age if the skill improvement channel is stronger that the entry cost channel, controlling for current density. Because this data set includes only entrepreneurs, it cannot be used to test the other implication of the learning model—that is, that all else equal, learning opportunities should increase the odds of selecting into entrepreneurship. The second data set we use (the Survey of Household Income and Wealth [SHIW] sample) addresses this question. The SHIW is a representative sample of the Italian population reporting, for each survey participant, the type of occupation, demographics (including place of birth and place of current residence), and data on personal income distinguished by source, such as income from entrepreneurial activity. However, it has no detailed information on the firm individuals work for or manage (besides size). Hence, the two data sets nicely complement each other. Finally, to investigate further and more directly which entrepreneurial abilities are learned through exposure to other firms earlier in life, we supplement the ANIA survey with measures of managerial practices collected using the methodology pioneered by Bloom and Van Reenen (2010b). If exposure to a larger set of firms allows one to learn superior managerial practices, then entrepreneurs who grew up in high firm density locations should adopt better managerial practices.

In our data sets there are two reasons why individuals living in a given province are exposed at learning age to different learning opportunities. First, they belong to different birth cohorts. Second, they grew up in different places. Hence, identification relies on two sources of variation: (*a*) differences over time in firm density for people belonging to different cohorts living in the same province where they grew up (stayers) and (*b*) cross-province differences in firm density for people belonging to the same cohort who grew up in a province different from the one in which they currently live (movers).[1] As we will show, focusing on the sample of movers addresses a series of endogeneity issues that can arise from the serial correlation between current entrepreneurial density (*ED* henceforth) and *ED* at learning age for entrepreneurs who did not move. In fact, our results are even stronger when focusing on the sample of movers. Finally, the two sources of variability also allow us to test for—and dismiss—the possibility that our *ED* indicator is proxying for other potential determinants of entrepreneurship, in particular differences in "culture" or the quality of education across locations.

Consistent with the learning model, we find that individuals who grew up in a location with a higher *ED* are more likely to become entrepreneurs. This result holds independently of whether we use a broader definition of entrepreneur (one that also includes the self-employed) or a narrower one that features only individuals running an incorporated business. Our finding holds while controlling for the firm density in the current location (reflecting thick-market externalities), for measures of current access to external finance in the local market where the firm is located and in the location at learning age, and for having parents who are entrepreneurs themselves. The effect is sizable. With the broader definition of entrepreneur, a 1 standard deviation increase in *ED* at learning age increases the likelihood of becoming an entrepreneur by 1.5 percentage points, around 8% of the sample mean. Not only overall density but also its sectoral distribution at learning age matters. Conditional on starting a business, the chances of starting it in a given sector increase with the sector density at learning age. The effects is large: a 1 standard deviation increase in a sector density at learning age raises the probability of being an entrepreneur in that sector by 10%. Because

[1] An Italian province is an administrative unit approximately equivalent to a US county.

sectors have idiosyncratic features and one is more likely to learn them if the sector is the one that dominates the local product market, this finding strengthens the interpretation that density causes entrepreneurship because it offers learning opportunities.

When we look at variation in performance among entrepreneurs, we find that those who faced a higher firm density at learning age earn a higher income from their business. A 1 standard deviation increase in firm density at learning age results in a 8% higher income. Because the SHIW reports only where a person was born and where he currently lives, this result is obtained under the assumption that an individual at learning age was located in the same place where he was born, thus inducing some measurement error in the firm density at learning age.[2] The ANIA sample is free from this problem, and in addition it allows construction of measures of firms productivity. In this sample we find that firms run by entrepreneurs who faced a higher firm density at learning age currently have a higher TFP and higher output per worker. Remarkably, the elasticity of entrepreneurial quality to *ED* is very close in the two data sets, despite the differences in sampling frame and time coverage.

The final question we address is which aspects of entrepreneurship are more prone to be learned. Classical theories of entrepreneurship stress the role of personal traits in terms of the ability to innovate (Schumpeter 1911) and to bear uncertainty and risk (Knight 1921; Kihlstrom and Laffont 1979). Modern literature on entrepreneurship argues that being an entrepreneur requires a variety of skills.[3] These features of entrepreneurship probably have an important innate component, and it is unclear to what extent they can be learned or significantly improved. On the other hand, managerial capabilities are potentially learnable. We therefore test whether entrepreneurs who grew up in high firm density provinces adopt better managerial practices and develop traits that are traditionally associated with entrepreneurship. We find some evidence that entrepreneurs who grew up in high firm density locations adopt better managerial practices. On the other hand, we find no evidence that exposure to firms at learning age affects the traits that have been traditionally associated with entrepreneurship, such as risk aversion, aversion to ambiguity, self-confidence, and optimism. These traits are either learned early in life, possibly within the family (Dohmen et al. 2012), or are truly innate.

Our paper contributes to several strands of literature. Closest to our work are studies of the effect of the environment when growing up on skill accumulation and occupational choice. Bell et al. (2018) show that growing up in

[2] This measurement error is likely to be small. Indeed, from the ANIA data set we calculate that 85% of people born in a given province are still in that province at learning age.

[3] For example, Lazear (2005) shows that MBAs with a more balanced set of skills (lower variance in exam grades) are more likely to become entrepreneurs.

an area with a high innovation rate in a specific technology class leads to a higher probability of becoming an inventor in exactly the same technology class. They interpret this as evidence of a causal effect of exposure to innovation on innovation propensity. Our case relates to a different outcome (the choice to become an entrepreneur and entrepreneurial success vs. being an inventor) in a different country (Italy vs. the United States). The fact that, despite these differences, the results are consistent across the two studies points to the importance of the environment in which one grows up in the determination of career choices and outcomes. De Figueiredo, Meyer-Doyle, and Rawley (2013) study "inherited agglomeration effects," defined as human capital that managers acquire while working in an industry hub that may be transferred to a spin-off. They focus on the hedge fund industry and show that hedge fund managers who previously worked in London or New York outperform those who did not in terms of financial returns on their portfolios. Like these papers, we also look at the effects of the environment during the learning age on subsequent labor market outcomes. We add to this work along two dimensions. First, we consider entrepreneurs. Second, by jointly analyzing the effects of learning opportunities on the propensity to become an entrepreneur and on performance as an entrepreneur, we are able to assess whether the dominant learning channel consists of a skill improvement effect or an entry cost reduction effect.[4]

Glaeser, Kerr, and Ponzetto (2010) document that area sectors with lower average initial firm size display higher employment growth in subsequent years. They show that the evidence is consistent with heterogeneity in both entry costs and in the supply of entrepreneurs. Our methodology to distinguish between learning to lower entry costs and learning entrepreneurial skills builds on Guiso and Schivardi (2011), who use the Lucas (1978) occupational choice model to set up a test to tell the two channels apart. Differently from these papers, we do not consider contemporaneous local characteristics but rather those prevailing at learning age.

We also contribute to the literature on the determinants of occupational choice. There is substantial evidence that the environment ("nurture") is at least as important as genetics ("nature") in the determination of the occupational choice, using the family as the channel through which nurturing takes place.[5] Our work contributes to this literature by showing that not only the

[4] Methodologically, this approach is related to the literature on wage city premiums. Glaeser and Maré (2001) show that a fraction of the urban wage premium—the extra wage that workers earn when moving to a city—stays with them when they move back to a suburban or rural area. They interpret this as evidence that workers accumulate human capital while in cities. This US-based evidence has been confirmed and extended by De La Roca and Puga (2017) for Spain and Matano and Naticchioni (2016) for Italy.

[5] See, e.g., Nicolaou et al. (2008), Zhang et al. (2009), Nicolaou and Shane (2010), and Lindquist, Sol, and Van Praag (2015).

family but also the local economic environment in which a person grows up affects the choice to become an entrepreneur.

The rest of the paper proceeds as follows. In section II we lay down a simple model of entrepreneurial choice, where agents learn how to avoid or minimize entry costs and how to acquire entrepreneurial skills. Higher learning opportunities shift to the right the initial distribution of entrepreneurial talent and soften entry barriers. This generates two testable predictions: higher learning opportunities (*a*) increase the chance that an individual becomes an entrepreneur and (*b*) raise the ability of those who select into entrepreneurship if the ability-shifting effect is sufficiently strong. We also discuss how we measure opportunities to learn entrepreneurial skills. In section III we discuss our identification strategy, while section IV presents the data. Results on occupational choice are shown in section V, while those on performance are shown in section VI. Section VII shows the evidence on managerial practices and traits, and section VIII concludes.

II. Learning, Entrepreneurial Skills, and Occupational Choice

In this section we provide a simple analytical framework to analyze the effects of heterogeneity in learning possibilities across locations and then discuss our measure of learning opportunities.

A. Modeling Learning Opportunities

We use the occupational choice model of Lucas (1978), as modified by Guiso and Schivardi (2011), to allow for multiple locations with different distributions of entrepreneurial skills and different entry costs. We illustrate the model briefly to derive some empirical predictions and refer the interested reader to Guiso and Schivardi (2011) for details. The economy is comprised of N locations, each with a unit population of workers who can choose to be an employee at the prevailing wage or become an entrepreneur. An entrepreneur combines capital and labor to produce output with a decreasing returns to scale technology and is the residual claimant. As such, entrepreneurial income is

$$\pi(x) = xg(k, l) - rk - wl - c, \tag{1}$$

where x represents entrepreneurial skills, k is capital, l is labor, r is the rental price of capital, w is the wage, and c is a fixed entry cost. The rental price and the wage are equalized across locations. As shown by Lucas (1978), the solution of the model can be characterized by a threshold value of entrepreneurial skills z such that an individual becomes an entrepreneur if and only if $x \geq z$. It is immediate to show that z depends on c and that $z'(c) > 0$: higher setup costs imply that the marginal entrepreneur has higher ability.

We assume that locations differ in terms of learning opportunities, parameterized by λ. Empirically, we will refer to λ in terms of different

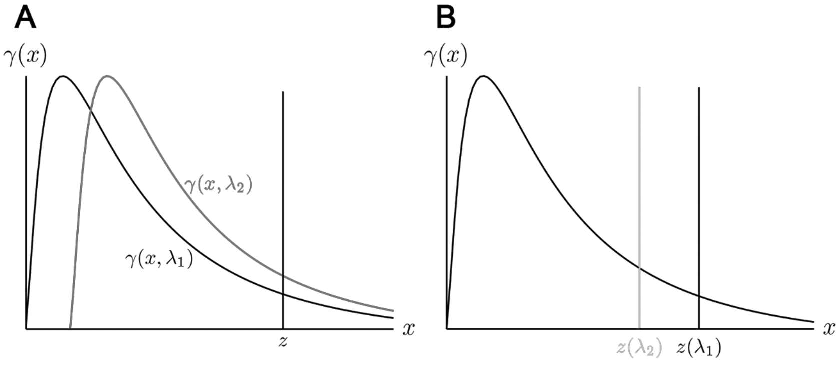

FIG. 1.—Shifts in skills distribution and entry costs. *A*, Ability distribution shift. *B*, Entry cost shift. A color version of this figure is available online.

measures of entrepreneurial density in the location in which an individual grows up. We let learning opportunities affect the occupational choice along two distinct dimensions. On one side, they allow individuals to accumulate more entrepreneurial skills. Growing up surrounded by entrepreneurs might allow an individual to observe them in action and therefore to learn how to successfully run a firm, for example, in organizing production, managing human resources, and dealing with suppliers and customers. We model this by assuming that entrepreneurial talent is a random variable X distributed according to a probability density function $\gamma(x, \lambda_j)$ over the support $(\underline{x}, \bar{x})$, $0 \leq \underline{x} < \bar{x} \leq \infty$, with corresponding cumulative density function $\Gamma(x, \lambda_j)$, $j = 1, \ldots, J$. The parameter λ is a shifter of the distribution of talent. It represents the learning opportunities that characterize each location.[6] We assume that $\partial\Gamma/\partial\lambda < 0$: λ shifts the probability distribution to the right in the first-order stochastic dominance sense. Hence, individuals who grow up in a region with high λ have on average higher entrepreneurial skills. Figure 1*A* plots the distribution of skills for two locations (1 and 2) with different learning opportunities, with location 2 having more opportunities ($\lambda_2 > \lambda_1$).

The second channel through which learning opportunities may affect the occupational choice is the setup cost c. We allow the cost to depend on λ, with $c'(\lambda) \leq 0$: learning opportunities can reduce the setup cost. Growing up in a location with a dense entrepreneurial structure might allow saving on entry costs, for example, because one learns how to obtain information on the bureaucratic procedures needed to set up a firm. Given that $z'(c) > 0$

[6] Of course, λ might be any shifter of the distribution of talent. Distinguishing learning from other possible explanations will be the main task of the empirical analysis.

and $c'(\lambda) < 0$, it follows that $dz/d\lambda = (\partial z/\partial c)(\partial c/\partial \lambda) < 0$: more learning opportunities reduce the ability threshold to became an entrepreneur. Figure 1*B* plots the ability threshold for two locations with different learning opportunities.

We now derive the effects of higher learning opportunities on the probability that an individual becomes an entrepreneur and on the average entrepreneurial skills of those who choose to become an entrepreneur. Given a threshold z, the probability of becoming an entrepreneur is $\Pr(x > z(\lambda)) = 1 - \Gamma(z(\lambda), \lambda)$. Taking the total differential, we obtain

$$\frac{d(1 - \Gamma(z, \lambda))}{d\lambda} = \underbrace{\left(-\frac{\partial \Gamma(z, \lambda)}{\partial \lambda}\right)}_{\text{skill improvement effect}} + \underbrace{\left(-\gamma(z, \lambda)\frac{\partial z}{\partial \lambda}\right)}_{\text{entry cost reduction effect}}. \tag{2}$$

Given the stochastic dominance assumption, the "skill improvement effect" is positive: in terms of figure 1*A*, the mass to the right of the threshold z is higher for the distribution parameterized by λ_2. Given that $\partial z/\partial \lambda < 0$, the "entry cost reduction effect" is also positive: in terms of figure 1*B*, there is more mass to the right of $z(\lambda_2)$ than of $z(\lambda_1)$. The two effects therefore reinforce each other and imply that higher learning opportunities lead to more entrepreneurs.

Predictions are slightly less obvious when it comes to average entrepreneurial skills. The average entrepreneurial skill level is the expected value of x conditional on being an entrepreneur:

$$\mathrm{E}(X|X \geq z, \lambda) = \frac{\int_{z(\lambda)}^{\bar{x}} x\gamma(x, \lambda)\, dx}{1 - \Gamma(z, \lambda)}. \tag{3}$$

The effect of a change in λ on average entrepreneurial quality is

$$\begin{aligned}\frac{d\mathrm{E}(X|X \geq z, \lambda)}{d\lambda} &= \underbrace{\frac{\left[\int_z^{\bar{x}} x \frac{\partial \gamma}{\partial \lambda}\, dx - \mathrm{E}(X|X \geq z, \lambda)\frac{\partial(1-\Gamma(z,\lambda))}{\partial \lambda}\right]}{(1 - \Gamma(z, \lambda))}}_{\text{skill improvement effect}} \\ &\quad + \underbrace{\frac{(\mathrm{E}(X|X \geq z, \lambda) - z)\gamma(z)}{1 - \Gamma(z, \lambda)}\frac{\partial z}{\partial \lambda}}_{\text{entry cost reduction effect}}.\end{aligned} \tag{4}$$

Let us first consider the entry cost reduction effect. When λ increases, the quality of the marginal entrepreneur decreases, and so does average ability. Again, this can be seen in figure 1, where it is clear that $\mathrm{E}(X|X \geq z(\lambda_1), \lambda) > \mathrm{E}(X|X \geq z(\lambda_2), \lambda)$.

The skill improvement effect cannot be signed a priori, as it depends on the distribution function of ability. However, the effect is positive for a general family of distributions: the log-concave distributions (Barlow and Proschan 1975).[7] This family of distributions includes, among others, the uniform, the normal, and the exponential. For such distributions, the skill improvement effect induces a positive correlation between the mass of entrepreneurs and their average quality. Hence, the total effect of learning opportunities on average ability depends on the relative strength of the two channels. If learning opportunities affect mostly the distribution of skills, then we should find that areas with more learning opportunities have both more and more capable entrepreneurs. If instead the entry cost effect dominates, areas with more learning opportunities will still have more but on average less skilled or talented entrepreneurs. Of course, the empirical challenge is to distinguish learning opportunities from other potentially correlated effects, given that we have no random variation in learning abilities. To this end, we will exploit mobility—that is, analyze individuals who grew up in a certain location and moved elsewhere as adults. As we argue below, this rules out the most obvious challenges to identification.

In the empirical analysis below we also study the role of sector-specific firm density. Suppose that learning opportunities are partly general and partly specific to a given sector. A simple intuitive extension of the model above is that—conditional on choosing to become an entrepreneur—an individual should be more likely to select the sector that has the highest density in the area where she grew up.

B. Entrepreneurs as Data Points

To test the implications of the simple model above we need an operational measure of λ—the opportunities to learn entrepreneurial skills. For this we assume that individuals growing up in different locations also face different learning environments because locations differ in the density of entrepreneurs active at a given point in time. Individuals who grow up in locations that are rich in firms (and entrepreneurs) have more opportunities to learn from the experiences of other entrepreneurs as part of their socialization process, compared with individuals who grow up in locations lacking entrepreneurs.

The idea that individuals acquire entrepreneurial capabilities from interacting and growing up among entrepreneurs is consistent with an expanding literature arguing that individual traits, besides being transmitted through parenting, are also acquired through socialization (Bisin and Verdier 2001), especially group and network contacts. Indeed, one strand of literature led by Hurrelmann (1988) and particularly Harris (2011) argues that interactions with peers dominate interactions with parents in the process of learning and personality formation. Furthermore, because group interactions develop

[7] A function $h(x)$ is said to be log-concave if its logarithm $\ln h(x)$ is concave, i.e., if $h''(x)h(x) - h'(x)^2 \leq 0$.

with age and become increasingly intense as young individuals start branching off from the restraints of their parents, these theories imply that the acquisition of entrepreneurial capabilities through social learning should peak when young individuals are in their late teens. Empirically, we will identify this age around 18 (which we label the "learning age") and proxy learning opportunities with the density of entrepreneurs in the area where individuals lived at their learning age.

Entrepreneurial density at learning age as a measure of learning opportunities captures the idea that it is easier to observe directly how entrepreneurs set up a firm (thus lowering entry costs) and their experiences with success and failures (improving managerial skills). For example, one needs to know how to obtain capital, how to identify a potential pool of investors, and how to "sell" the idea to them. This capability can be learned, possibly from other entrepreneurs. This makes business creation easier. Observing other entrepreneurs also offers insights on what leads to success as well as how to avoid mistakes that lead to failure, allowing selection of better ideas and to implement them more effectively. This mechanism is consistent with Chinitz (1961), who first documented that entry of new firms is more likely where a high number of small businesses is present. Our empirical strategy is also consistent with the plausible idea that there are stages of learning through socialization characterized by the different content of what is learned. For our purpose, the formative years—those that define what an individual would like to be and what she can become as an adult—belong in the 18-year-old age bracket (Erikson 1968). Feedback from other entrepreneurs concerning the content of their work, the requirements to succeed in doing it, the type of life one can expect from selecting an entrepreneurial job, "role modeling," and so on can be critical at this age. How important it is may depend on the number of learning points an individual is exposed to.

Needless to say, learning may occur not only from firms (before one starts to work) but also within firms (as worker, apprentice, etc., at some of the local firms). Thus, while we take age 18 as a "reference" learning age, entrepreneurial learning is likely to start at an earlier age and continue as people advance in their early adulthood. In fact, because entrepreneurial density evolves very slowly over time, the measure of entrepreneurial density at age 18 should be thought of as a proxy for entrepreneurial density when young more generally. We provide evidence consistent with this view below.

It is worth noting that the agglomeration literature has stressed two main sources of entrepreneurial learning. On one side, there are the so-called "urban" externalities, typically associated with Jacobs (1969), who stressed the role of having access to a diverse set of learning opportunities. On the other, there are the "specialization" externalities, put forward by Marshall (1890), who pointed out that learning is specific to a sectoral activity. In our setting, this translates into learning occurring from all entrepreneurs or just from entrepreneurs in the same sector in which an individual chooses to operate.

We take no a priori stand on this, as there are arguments that make both plausible, and let the data speak. We first use the overall entrepreneurial density—that is, the total number of entrepreneurs over the population—as the measure of learning opportunities. This is the natural starting point to analyze the occupational choice: entrepreneurs versus paid employment. We then study the effect of the sectoral density on the choice of the sector in which to operate, conditional on being an entrepreneur. If sector-specific learning opportunities matter, either because one acquires sector-specific skills or learns about sector-specific set up costs, this should be detectable first of all in the choice of sector.

III. Identifying Learning Effects

Our empirical strategy is based on two broad sets of regressions. First, we determine whether growing up in an area with higher entrepreneurial learning opportunities (as measured by the number of firms per capita, the variable entrepreneurial density [*ED*] defined below) is associated with a higher likelihood of becoming an entrepreneur; conditional on becoming an entrepreneur, we also test whether sectoral density affects sectoral choice. Occupational and sectoral choice cannot tell whether density matters because it improves entrepreneurial skills or because it lowers entry costs, as both effects act in the same direction. We then turn to studying measures of entrepreneurial performance. As predicted by the model, this second type of analysis identifies whether the dominant learning channel is skill enhancement or reduction in entry costs. We now discuss the main empirical challenges we face in bringing these predictions to the data. The discussion focuses on overall entrepreneurial density; at the end of the section, we elaborate on the additional identifying power deriving from the sectoral component of entrepreneurial density.

Our empirical framework is based on regressions of the form[8]

$$Y^*_{it} = \alpha + \beta ED_{j(i,t_L)t_L} + \gamma ED_{j(i,t)t} + \varepsilon_{it}, \tag{5}$$

where Y^* is either the net utility from being an entrepreneur or a measure of entrepreneurial performance, ED is firm density (number of firms per capita), and ε is an error term. The subscripts $j(i,t_L)$ and $j(i,t)$ represent the location in which individual i was living in year t_L and year t, respectively, with $t_L < t$ being the year in which individual i turned 18 (the learning age) and t being the current year. Since we do not observe the utility from being an entrepreneur, we define the entrepreneurship indicator $Y_{it} = 1\{Y^*_{it} > 0\}$. Under the assumption that ε_{it} is standard normally distributed, we can thus estimate probit models for the decision to be an entrepreneur.

[8] For notational simplicity we omit the vector of additional controls used in all regressions. These are discussed later.

The role of current firm density ($ED_{j(i,t)t)}$) in these regressions is well known from the literature on local externalities and agglomeration economies.[9] The role of firm density at learning age ($ED_{j(i,t_L)t_L}$) is instead the channel we emphasize in this paper—learning entrepreneurial skills in the early phase of one's professional life—and in our framework it may be present over and above that of $ED_{j(i,t)t}$. Naturally, an empirical challenge is that the effect of $ED_{j(i,t_L)t_L}$ may be hard to identify separately from that of $ED_{j(i,t)t}$ given the persistence in the spatial agglomeration of firms. There are two distinctive sources of variation that allow us to separately identify the two effects. First, young-age learning externalities can be distant in the time dimension from current externalities, which is of course especially true for older entrepreneurs and if geographical locations go through phases of industrial booms and decays. Second, some individuals currently live in locations that are different from those in which they grew up (movers), spatially breaking the link between $ED_{j(i,t_L)t_L}$ and $ED_{j(i,t)t}$. In addition to using the overall sample, we will also run regressions on the sample of movers to provide a more compelling identification of the effect of entrepreneurial learning opportunities on entrepreneurial outcomes. Note that since we are controlling for current entrepreneurial density, identification in the whole sample comes from changes in entrepreneurial density between the year an individual turns 18 and the current year of observation. Differences in the growth path may be correlated with other determinants of becoming an entrepreneur and bias the estimates. We now discuss these issues in detail.

A natural concern is omitted variable bias: there could be unobserved factors that determine both an entrepreneur's success (or the net benefits of being one) and the entrepreneurial density in a given area. For example, a well-functioning local financial system might be able to lower entry barriers and at the same time screen the best entrepreneurial projects, inducing an upward bias in our estimation of γ in equation (5). Alternatively, suppose that there are government programs that subsidize entrepreneurship or provide management training in low firm density areas, thus decreasing the costs of being an entrepreneur as well as improving business performance. This would induce a downward bias in γ. Given that $ED_{j(i,t)t}$ is strongly serially correlated, any bias induced by these omitted variables will also transmit to β, our effect of interest. Ideally, an instrument would address this concern. Unfortunately, finding instruments for entrepreneurial density (and for agglomeration economies in general) is difficult and is particularly so in our case, where a time-varying instrument would be needed. We therefore adopt two empirical strategies to address this issue. The first is a traditional one: control for a rich set of characteristics (including demographics, geographical controls, intergenerational variables, and controls for both current

[9] For surveys of this literature, see Duranton and Puga (2004), Rosenthal and Strange (2004), and Moretti (2011).

and past local credit market development), which in principle minimizes the set of unobservables that may potentially be correlated with current firm density. The second is to use the sample of movers. If provinces had purely idiosyncratic dynamics in firm density, the endogeneity bias induced by the correlation between $ED_{j(i,t)t}$ and ε_{it} would not "transmit" to the effect of $ED_{j(i,t_L)t_L}$ onto Y^*_{it} in a sample of movers (because $j(i, t_L) \neq j(i, t)$). However, as we document below, provinces do have a common time dynamics induced by overall economic growth. In other words (simplifying the subscript notation slightly), firm density appears to follow an AR(1) process with an aggregate drift:

$$ED_{jt} = \mu_t + \rho ED_{jt-1} + \zeta_{jt}, \tag{6}$$

where it is assumed that $E(\zeta_{jt}\zeta_{ks}) = 0$ for all $\{j \neq k\}$ and $\{s \neq t\}$. For stayers (for whom $j(i, t) = j(i, t_L) = j$), we are regressing

$$Y^*_{it} = \alpha + \beta ED_{jt_l} + \gamma ED_{jt} + \varepsilon_{it},$$

and, given equation (6), ED_{jt_L} will naturally be correlated with ED_{jt}, implying that if $E(\varepsilon_{it}|ED_{jt}) \neq 0$ then $E(\varepsilon_{it}|ED_{jt_L}) \neq 0$ as well. For movers (for whom $j = j(i, t) \neq j(i, t_L) = k$), however, we will be regressing

$$Y^*_{it} = \alpha + \beta ED_{kt_L} + \gamma ED_{jt} + \varepsilon_{it},$$

and ED_{kt_L} and ED_{jt} are correlated only because of aggregate effects (especially recent ones, as farther ones exert diminishing effect as long as $\rho < 1$). In other words,

$$\mathrm{E}\left(\varepsilon_{it}|ED_{jt}, \mu_t, \mu_{t-1}, \dots\right) \neq 0$$

but

$$\mathrm{E}(\varepsilon_{it}|ED_{kt_L}, \mu_t, \mu_{t-1}, \dots) = 0.$$

Controlling for year effects, therefore, breaks the correlation between current firm density and firm density at learning age in the movers sample. Since our interest centers on the identification of β, this is enough to obtain unbiased estimates of β.[10]

Using movers, however, may introduce a sample selection (sorting) bias, since the choice of moving from $j(i,t_L)$ to $j(i,t)$ may have been driven by unobservables correlated with entrepreneurship: for example, individuals with a good business idea may move to areas with a higher firm density because these are more attractive locations to start a business. We address this

[10] If firm density is subject to spatially correlated shocks (e.g., within a cluster of geographical locations) besides aggregate shocks, controlling for time effects would not be enough for people who move within a cluster. For robustness, we also consider specifications in which we drop within-cluster movers.

form of sorting using a Heckman probit selection model for location mobility and use aggregate migration rates as an exclusion restriction (see below for details). A different form of sorting is with respect to $ED_{j(i,t_L)t_L}$—that is, people move to areas where opportunities for entrepreneurial learning are better. While this is possible in principle, in practice we do not believe it is a serious concern. First, the location where an individual grows up is (arguably) chosen by the parents, and it seems quite unlikely that parents locate in specific areas to indulge their offspring's entrepreneurial attitudes (which may be hardly detectable at a young age to begin with).[11] Still, one could argue that parents with entrepreneurial traits choose to locate in entrepreneurially dense areas and that such attitudes are genetically or culturally transmitted to the offspring. Our control for having a parent entrepreneur captures a substantial part of these traits but, possibly, not all: some parents with these traits may still move to dense areas but not become entrepreneurs. To account for this, in a robustness check we enrich the set of parental controls, including education and sector of activity.

Finally, unlike the one discussed above, a different bias may come from omitted variables that determine both *ED* at learning age and persistently affect the propensity to become an entrepreneur and entrepreneurial ability even after one moves from the learning age location. The presence of these variables represents a threat to identification even in the sample of movers. For example, differences in culture or in the quality of the school system might both determine heterogeneity across locations in *ED* at learning age and individual entrepreneurial outcomes later in life. Following Max Weber's culture theory, in certain areas entrepreneurship might be regarded as a particularly appealing occupational choice and entrepreneurial success as a highly regarded outcome. Because culture is persistent, an individual who grows up in a high *ED* area might both be more likely to become an entrepreneur and exert more effort in entrepreneurship as a reflection of the culture of the place where she grew up. This story would give rise to the same type of correlation implied by the learning story and, because culture is portable, would also affect the specification where we focus on movers.

To address these threats to identification in our movers sample, we use the argument that learning entrepreneurial abilities from entrepreneurial density presumably evolves at different frequencies and geographical reach than culture. In fact, we will show that *ED* changes substantially over the sample period even after netting out aggregate time effects, giving rise to nontrivial within-location time series variation. Culture is instead a process that is likely to move at very low frequencies (Williamson 2000). Furthermore, while learning is local (you learn from people you directly interact with), culture

[11] In one of our robustness checks we drop individuals who moved between birth and age 18, eliminating this concern at its source.

typically spans broader geographical areas. Hence, it can be accounted for by broader geographical controls than those that define variation in entrepreneurial density. Stated differently, if fixed local attributes are important, then $E(\varepsilon_{it}|ED_{j(i,t_L)t_L}) \neq 0$, while $E(\varepsilon_{it}|ED_{j(i,t_L)t_L}, Geo_s) = 0$, where Geo_s are detailed geographical dummies for both the current and the learning age location. Therefore, by comparing our estimates as we vary the number of spatial dummies (making them finer), we are able to assess the likelihood that fixed local attributes represent a credible threat to identification.

The other potential confounding factor is heterogeneous school quality. If a given geographical area is endowed with better schools and school quality is a determinant of entrepreneurship, it will produce more and better entrepreneurs. Moreover, the effect will be long lasting, as schooling gets embedded in the human capital that travels with the individuals on moving. We address this concern directly by documenting that there is no correlation between school quality and entrepreneurial density.

The discussion so far has focused on overall entrepreneurial density. Bell et al. (2018) argue that the sectoral component of the entrepreneurial choice offers additional identifying power for the causal effect of learning opportunities on the propensity to engage in entrepreneurship. Intuitively, if we find that an individual who grew up in Prato (an area with a dense concentration of textile firms) is more likely to run a textile business—conditional on being an entrepreneur—even if she moves to Parma (a food district), this would be evidence that exposure at learning age has a causal effect on the sectoral choice. This reasoning holds under the realistic assumption that those who grow up in Prato do not have a higher innate propensity to start a textile business relative to other types of businesses.

IV. Data

A. The SHIW Sample

SHIW collects information on demographics, income, and assets for a representative sample of Italian households. Starting in 1991, the survey is run biannually (with the exception of 1997), and we use all 11 waves from 1991 to 2012 for a total of 62,756 observations (all household members who are employed and aged 30–65). For our purposes, the SHIW contains data on occupations and earnings from various sources—including earnings from business—for each household member. Moreover, for each individual it reports the province of birth and the province of residence. The province, an administrative unit comparable in size to a US county, is our geographical reference for measuring learning opportunities in the SHIW sample. There are 95 provinces at the start of the sample period.[12] To identify entrepreneurs,

[12] Over the sample period new provinces were created by split off of existing provinces; we use the initial 95 province classification.

we use two measures. The first is a broad measure that includes people who are self-employed, partners of a company, and owners who run an incorporated business (19% in total). The second is a narrow definition that includes only the latter category (9% of the sample); it replicates the entrepreneur definition in the ANIA survey described next. This sample allows us to study the occupational choice but has limited information on the firm's characteristics and the sector of activity, recorded at the one-digit level (four sectors only). Panel A of table 1 shows summary statistics for the SHIW sample. The key variables are defined in the appendix.

B. The ANIA Sample

Our second data source consists of detailed information from a sample of entrepreneurs and their firms. The data set is based on a survey conducted by ANIA, covering 2,295 private Italian firms employing between 10 and 250 employees. The survey was conducted between October 2008 and June 2009. It consisted of two distinct questionnaires. The first collected general information on the firm and was filled out by the firm officials on a paper form. The focus of this first questionnaire was on the type of firm-related insurance contracts that the firm had or was considering. The questionnaire also collected more general information about the firm (such as ownership structure, size, and current performance) and its demographic characteristics. The second questionnaire collected information on the person in charge of running the firm. This questionnaire was completed in face-to-face interviews by a professional interviewer using the computer-assisted personal interviewing (CAPI) method. Several categories of data were collected, including information on personal traits and preferences, individual or family wealth holdings, family background, and demographics. The latter in particular includes information on the municipality where the individual (as well as his or her spouse) was born and where he or she was living at 18.

For approximately half of the firms (those incorporated as limited liability companies), we also have access to balance sheets. These data were provided by the Cerved Group, a business information agency operating in Italy. The data from the two sources were matched using a uniquely identifying ID number. The matched sample is the one we use in this paper. The data necessary to compute TFP are available for the years 2005–7. We end up with 966 firms and almost 2,600 firm-year observations. TFP is computed using factor shares, assuming constant returns to scale (results are robust to alternative computation methods). Different from the SHIW sample, we know the municipality (a lower administrative unit than the province) in which the business is located. As our preferred geographical unit we thus use the local labor systems (LLS), that is, territorial groupings of municipalities characterized by a certain degree of working-day commuting by the resident population, which represent self-contained labor markets and are

Table 1
Descriptive Statistics

Variable	Mean	SD	Variable	Mean	SD
A. SHIW Sample:			B. ANIA Sample		
Entrepreneur, definition 1	.19	.39	TFP (log)	2.45	.90
Entrepreneur, definition 2	.09	.28	*ED* learn	.06	.019
ED learn	.05	.02	*ED* today	.09	.015
ED today	.08	.02	ED*SET* learn	.004	.007
Male	.62	.49	ED*SET* today	.007	.007
Age	44.60	8.51	Male	.67	.47
Experience	24.32	10.25	Age	46.03	10.47
Parent entrepreneur	.13	.33	Experience	15.84	11.00
Elementary school	.12	.33	Parent entrepreneur	.36	.48
Junior high school	.33	.47	Elementary school	.01	.08
High school degree	.39	.49	Junior high school	.07	.25
College	.14	.35	High school degree	.68	.47
Postgraduate	.01	.08	College	.23	.42
Married	.77	.42	Postgraduate	.02	.13
Family size	3.38	1.18	Married	.78	.41
Number of income recipients	2.01	.82	Family size	3.06	1.17
Mover	.22	.42	Mover	.26	.44
Mover-entrepreneur	.19	.39	Learn in northwest	.30	.46
Income (log)	9.04	.87	Learn in Northeast	.29	.45
Born northwest	.19	.39	Learn in center	.19	.39
Born northeast	.21	.40	Learn in south	.22	.41
Born center	.20	.40	Resident in northwest	.31	.46
Born south	.40	.49	Resident in northeast	.30	.46
Resident northwest	.24	.43	Resident in center	.20	.40
Resident northeast	.22	.41	Resident in south	.19	.39
Resident center	.22	.41	Mining	.03	.17
Resident south	.32	.47	Manufacturing	.35	.48
Agriculture	.05	.21	Utilities	.01	.09
Manufacturing	.23	.42	Construction	.08	.27
Construction	.07	.25	Trade	.24	.43
Services and others	.66	.47	Transport	.04	.20
Firm size	1.11	18.78	Other services	.24	.43
LC learn	.22	.11	Employees	34.30	40.38
LC Today	.51	.20	LC learn	.25	.11
GDP growth	.05	.04	LC Today	.62	.17
			GDP growth	.04	.03

NOTE.—See sec. A4 of the appendix for definitions of variables. ANIA = Associazione Nazionale delle Imprese Assicurative; GDP = gross domestic product; SHIW = Survey of Household Income and Wealth; TFP = total factor productivity.

therefore the ideal geographical unit within which to study local externalities. LLS are similar to US metropolitan statistical areas. We use the definition based on the 2001 census, which identifies 686 LLS.[13] Results are robust when performing the analysis at the provincial level.

[13] In contrast, there are only 95 provinces.

Summary statistics for this sample are shown in panel B of table 1. The comparison with the SHIW sample indicates that they are fairly similar. The main differences are that the ANIA entrepreneurs are on average more educated, are less likely to have grown up or be resident in the south, and manage larger firms. Note that some of the differences are expected, since the SHIW statistics refer to both entrepreneurs and nonentrepreneurs, while ANIA is a sample of entrepreneurs. Moreover, the ANIA sampling scheme excludes firms with less than 10 employees and only includes incorporated businesses. The appendix describes the design of the SHIW and ANIA surveys in greater detail and provides a precise description of the variables used in this study.

The two data sets we use have advantages and disadvantages. The SHIW sample is representative of the Italian population and as such is ideal to study the decision to become an entrepreneur. The ANIA survey is instead a sample of entrepreneurs. On the other hand, the SHIW reports only the current place of residence and the residence at birth, while the ANIA also contains the location around age 18. Moreover, because the ANIA data are relative to firms, they allow us to construct the direct empirical counterpart of the Lucas model's measure of entrepreneurial ability, that is, TFP.

C. Measuring Learning Opportunities and Other Controls

We measure learning opportunities with firm density at learning age. As explained above, as a reference measure of location we use provinces for the SHIW sample and LLS for the ANIA sample.[14] To measure firm density we obtain census data on both the population and the number of firms active in each location and year since 1951 and divide it by the corresponding resident population.[15] We then attach to each individual in our sample (in either the SHIW or the ANIA survey) the firm density at the location at learning age. Because the census data are available for 1951, 1961, 1971, 1981, 1991, 2001, and 2011, we perform a simple linear interpolation at the province or LLS level for the midcensus years.

We follow a similar procedure to construct entrepreneurial density at the sectoral level. This task is complicated by the substantial changes in the sectoral classification that occurred over the six decennial censuses we use. The finest homogeneous sectoral classification we were able to reconstruct across censuses is based on 33 sectors. The appendix details the sectoral concordance procedure.

[14] This is due to data restrictions, as the province is the SHIW's lowest available level of geographical disaggregation.

[15] In the choice of the firm definition we are constrained by data availability in the early censuses. In particular, they only report data for production units, which are similar to a plant. Moreover, they do not distinguish between business units and other types, such as government units. We therefore keep all units throughout. In the most recent censuses, nonbusiness units account for less than 3% of the observations.

In the ANIA sample we know where each entrepreneur was living at learning age and can attach ED in the location where she was at that age. In the SHIW we know where the individual was born but not where she grew up; in these cases we assume that they grew up in the province of birth. Hence in this sample our measure of learning opportunities contains some measurement error. This measurement error, however, is likely to be small. From the ANIA sample (where we observe both the place of birth and the place at learning age) we calculate that only 15% of people grew up in a province different from that of birth. In SHIW we can calculate patterns of mobility between birthplace and current location. The transition matrix, reported in the online appendix (table OA.1), shows much larger mobility rates from the south and islands (79% stay, 16% migrate to the north, 5% migrate to the center) than in the center (94% stay, 5% migrate to the north, 1% migrate to the south) and the north (97% stay, 2% migrate to the center, 1% migrate to the south).

Our identification exploits both cross-sectional and time series variation in firm density at learning age. Figure 2*A* shows the pattern of firm density over time for each province in the sample, while panel B focuses on the largest Italian provinces.[16] Density differs considerably across provinces at each point in time as well as over time within provinces with very different time profiles. Consider individuals currently living in a given province X. They differ along two dimensions: their current age and the province where they grew up. Some grew up in the same province where they currently live, while others moved after spending their formative years in a different province. We can identify the effect of learning opportunities through two thought experiments. Everything else equal, we can compare the occupational decision and the performance as entrepreneurs of individuals currently located in province X who grew up in X in different time periods and hence faced different firm densities at learning age (i.e., individuals who grew up in Florence in the early 1950s vs. the early 1990s). Or we can compare the outcomes of individuals of the same age who grew up in different provinces and hence faced different firm densities at learning age before moving to province X (e.g., individuals who are currently in Florence but grew up in different provinces, say Milan or Palermo). Because firm density is persistent, current density and density at learning age for individuals based in the province where they grew up tend to be relatively highly correlated (correlation coefficient, 0.6), especially for younger individuals. Identification is facilitated by movers. For this subsample of individuals (22% in the SHIW and 26% in the ANIA sample), correlation between current density and density at learning age is much lower (0.2). In terms of the relative importance of the two types of

[16] In fig. A1, we show the spatial distribution of firm density for two census years, 1951 and 2011. The comparison shows substantial changes in the geographical distribution of density between the two censuses.

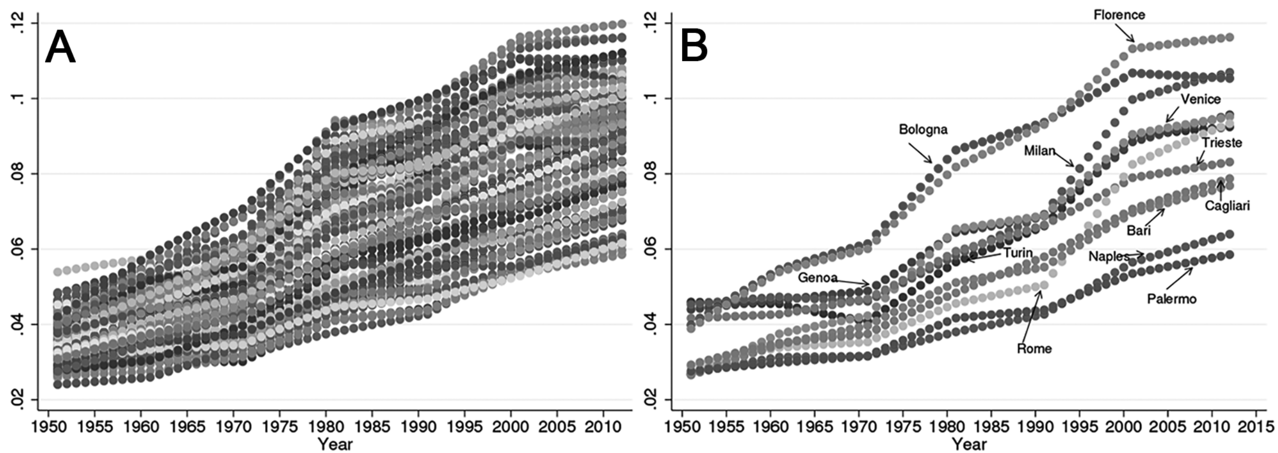

FIG. 2.—Evolution of firm density across Italian provinces, 1951–2009. *A*, All provinces. *B*, Largest provinces. A color version of this figure is available online.

variations to overall variation in density, the variance decomposition shows that it is rather similar: 52% is accounted for by the between component, and 48% is accounted for by the within component.

Finally, we complement our data sets with the number of bank branches per capita in each location-year as a measure of local credit market development.

V. Occupational and Sectoral Choice

We begin by studying the occupational choice. We do it in two steps. First, using the SHIW sample, we ask whether entrepreneurial density at learning age increases the likelihood that an individual chooses entrepreneurship over paid employment. Second, given that the SHIW data set reports the sector of occupation only at the one-digit level, we use the ANIA sample to analyze the relationship between sectoral entrepreneurial density at learning age and the sectoral choice of entrepreneurs. Importantly, we study these effects both on the overall sample and restricting the analysis to the sample of movers. As argued in section III, this offers further support to the causal interpretation of density at learning age on subsequent entrepreneurial choices and outcomes.

A. SHIW Sample: Occupational Choice

We start by estimating a probit model for the binary decision to become an entrepreneur. In our specifications, besides firm density in the province where the individual was located at learning age, $ED_{j(i,t_L)t_L}$ (here assumed to be the same as the province of birth), we control for firm density in the province of current residence, $ED_{j(i,t)t}$. We capture general geographical features that may affect occupational decisions (such as the cost of starting a business) by inserting dummies for the area of birth and for the current area of residence of the individual (either four macro area dummies—northeast, northwest, center, and south—or 20 regional dummies). In addition, we control for individual demographics, such as gender, age, educational attainment, work experience, a dummy for whether the parents were entrepreneurs, and family characteristics (whether married, number of earners, and family size). A key issue in the choice to become an entrepreneur is access to finance. A large literature argues that liquidity constraints and easiness in raising external capital foster entrepreneurship (see, e.g., Evans and Jovanovic 1989; Banerjee and Newman 1994). It might be that firm density at learning age reflects local financial development, as more firms may be started where capital is easier to raise. To account for this we control for the number of bank branches per capita in the province at learning age and for the same variable in the province of residence at the time the survey was run. As shown by Guiso, Sapienza, and Zingales (2004), the number of bank branches per capita predicts the easiness in obtaining external finance. To capture cohort-specific growth opportunities that may affect the choice of entrepreneurship as well as subsequent performance, in all regressions we also control for regional gross domestic product (GDP) growth at learning age. All regressions contain (unreported) dummies for education, year, and sector. Given that our main variable of interest (*ED* at learning age) varies at the province-year level, we cluster standard errors accordingly. Results are robust to alternative clustering schemes.

Results are shown in table 2 for the broader definition of entrepreneurship, which includes the self-employed, partners of a company, and owners who run an incorporated business. Marginal effects are reported throughout. In column 1 we report the result of a regression with *ED* at learning age without controlling for current *ED*. People who grew up in provinces with a higher firm density are more likely to become entrepreneurs. As for the other controls, bank branches per capita at learning age have no statistically significant effect. Males are more likely to be entrepreneurs, as are older and married individuals. Having a parent who was an entrepreneur has a strong positive impact on the likelihood of being an entrepreneur, consistent with most of the literature on the determinants of entrepreneurship.[17] Moreover, the number of income recipients within a household also exerts a positive effect, arguably because employed family members are a source of startup capital and income insurance, which may be important to smooth out entrepreneurial income fluctuations.

In column 2 we add current *ED*. This is a key control given that for stayers current and learning age *ED* are correlated, so that the latter might just be proxying for the former. We also control for current bank branches per capita. As expected, the coefficient on learning age *ED* decreases, from 1.17 to 0.73, but remains large and statistically significant. Current *ED* has a slightly larger coefficient (1.1) and is also significant.[18] Increasing *ED* at learning age by 1 standard deviation increases the probability that an individual decides to become an entrepreneur by 1.5 percentage points, 8% of the sample mean.

[17] In a series of unreported exercises, we have added dummies for education and for the sector of activity of the parents to control for the possibility that parents with more entrepreneurial traits choose to locate in denser areas in terms of entrepreneurship, as discussed in sec. III. Results are virtually identical.

[18] Note that while the correlation between the choice to become an entrepreneur and current *ED* may suffer from a form of reflection problem (Manski 1994), our variable of interest—*ED* at learning age—is immune from it. Of course, current *ED* may be potentially endogenous, as both this variable and individual behavior may respond to simultaneous shocks. The bias in the estimates might transmit to other variables, particularly in nonlinear models. This concern is alleviated by the fact that the individuals in our sample decided to become entrepreneurs on average 15 years ago (with a standard deviation of 11), which implies that their occupational choice may be independent of current shocks. Moreover, in what follows we provide robustness checks along many dimensions, including the geographic ones, finding that the estimates are remarkably stable.

Table 2
Probability of Becoming an Entrepreneur, Survey of Household Income and Wealth (SHIW) Sample, Definition 1

	(1)	(2)	(3)	(4)
ED learn	1.170***	.734**	.646*	3.535***
	(.282)	(.321)	(.373)	(1.164)
ED today		1.092***	2.087***	.176
		(.333)	(.438)	(1.184)
LC learn	−.006	−.002	−.000	−.012
	(.008)	(.009)	(.011)	(.021)
LC today		−.024	−.016	−.141***
		(.017)	(.020)	(.044)
Male	.043***	.043***	.043***	.066***
	(.005)	(.005)	(.005)	(.016)
Age	.142***	.127***	.123***	.197***
	(.022)	(.023)	(.025)	(.061)
Experience	−.020***	−.020***	−.019***	−.029
	(.006)	(.006)	(.006)	(.018)
Parent entrepreneur	.136***	.137***	.136***	.098***
	(.007)	(.007)	(.007)	(.019)
Married	.019***	.018***	.018***	.006
	(.006)	(.006)	(.006)	(.021)
Family size	.001	.001	.001	.002
	(.002)	(.002)	(.002)	(.006)
GDP growth at learning age	−.034	−.040	−.031	.027
	(.068)	(.068)	(.067)	(.175)
Number of income recipients	.012***	.011***	.011***	.018**
	(.003)	(.003)	(.003)	(.007)
λ_{mover}				−.638***
				(.180)
Observations	62,756	62,756	62,756	13,360
Area dummies:				
Macro area of birth	X	X		X
Macro area of residence	X	X		X
Region of birth			X	
Region of residence			X	

NOTE.—Shown are results of probit regressions for the choice of being an entrepreneur, marginal effects. Entrepreneur definition 1 includes (*a*) individual entrepreneurs, (*b*) owner or member of family business, (*c*) working shareholder/partner, and (*d*) self-employed/craft workers. *ED* learn is entrepreneurial density at 18 in the place of birth, *ED* today is current entrepreneurial density in the place of residence. LC learn and LC today are indicators for liquidity constraints at learning age and in survey year, respectively. All regressions include year, education, and sector dummies. Column 4 uses only the sample of movers, correcting for selection with a Heckman model in which the excluded variables are the average mobility rate out of the region of birth in the 10 years before and after the learning age. Standard errors clustered at the level of year of learning and province are in parenthesis. GDP = gross domestic product.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Having established this basic pattern, we now check whether it is robust to a number of potential objections. As discussed in section III, there is still the possibility that our results are driven by unobserved local factors that drive both *ED* at learning age and entrepreneurial outcomes later in life, such as school quality and differences in culture. Regarding school quality, ideally one would like to control for it at the local level at the time of learning. Unfortunately there is no source of information on school quality at the local level over the required time span and thus we cannot control for this potential confounding factor in the regressions. However, we do observe school quality in recent years. If, as the objection holds, school quality is higher in high *ED* areas and this leads to higher human capital and possibly to more or better entrepreneurs, we should find a positive correlation between school quality and *ED* even today. Since 2008, Italian fifth graders have taken a national standardized test in reading and math conducted by the Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e Formazione (INVALSI). The test, similar to the Organization for Economic Cooperation and Development's Program for International Student Assessment (PISA) test, is administered to more than half a million students in 6,000 schools in 3,400 cities. We average the 2008–9 INVALSI test scores measure at the province level and regress this variable on current *ED*, after controlling for macro area dummies. Figure 3 reports the regression line for test scores in language and math. We find no correlation for language scores and a slightly negative correlation for math scores. This evidence goes against the objection that our measure of *ED* may be proxying for unobserved school quality.

To address the other potential source of unobserved heterogeneity—cultural differences across areas—we increase the number of spatial controls, reducing the contribution of the cross-sectional variability in the data to identify the parameters and thus exploiting mostly the time series variation. The assumption is that such alternative determinants tend to move at much lower frequencies than *ED* (and learning entrepreneurial skills from it). Moreover, as argued above, learning is likely to be more localized than culture, whose effect should be accounted for by finer geographical dummies. Both arguments imply that the coefficient of *ED* should become smaller as we increase the number of spatial controls if *ED* is mostly capturing cross-sectional differences in cultural endowments but should remain roughly unchanged if *ED* at learning age captures the possibility of accumulating entrepreneurial ability. In column 3 we introduce 20 regional dummies, both for the current location and for the area of learning location.[19] Results for the *ED* at learning age variable are similar, suggesting that fixed unobserved local characteristics are unlikely to be driving our estimates.

Finally, in the last column we focus on the sample of movers, that is, individuals who were born in a province different from their current province

[19] Regions are territorial units composed on average of five provinces. We have also experimented with province dummies, which completely eliminate the cross-sectional differences. The point estimates are similar, but we lose statistical precision.

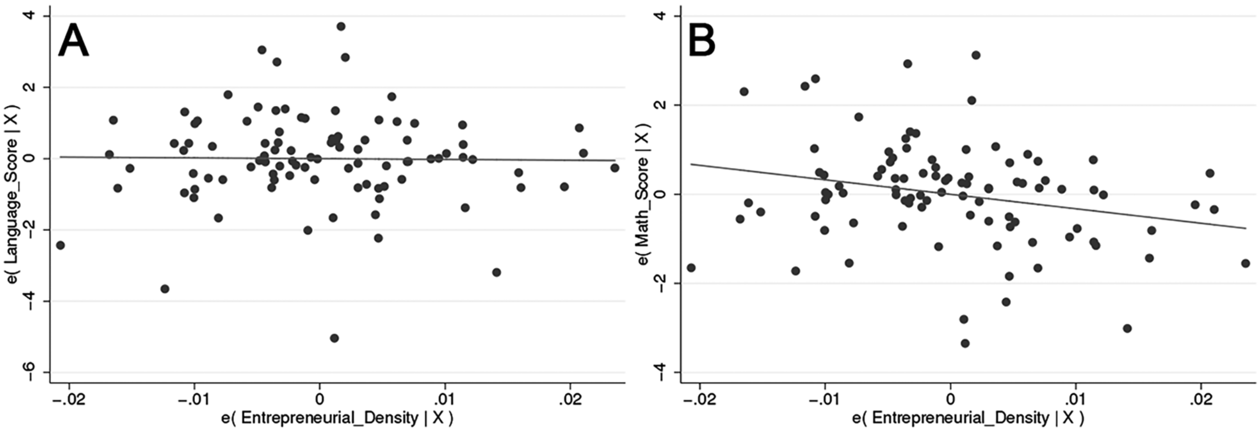

FIG. 3.—School test scores and entrepreneurial density, Italian provinces. *A*, Language score. *B*, Math score. A color version of this figure is available online.

of residence. This comparison allows us to break the collinearity between current and learning age density. To control for selection into moving, we run a Heckman sample selection model. As an exclusion restriction for mobility we use data on internal migration flows from the National Institute of Statistics. Given that our geographical unit is the province, ideally we would like to have data on out-of-province migration. Unfortunately, the data only contain (for each region and year) information on the number of individuals who move out of the region (a larger administrative unit than the province) and the (aggregate) number of individuals moving out of a municipality (a smaller administrative unit than the province).[20] The latter is a combination of intra- and interregional mobility. Since provinces are closer to a region than to a municipality (there are 20 regions, 95 provinces, and 8,100 municipalities in Italy), we focus on out-of-region migration rates, which we obtain by dividing the annual number of out-of-region migrants by the regional population. Given that we do not know the exact year of the move (as we only know province of birth and the current province of residence), we take the average mobility rate in the 10 years before and the 10 years after learning age and include them in the first stage of the Heckman procedure.

We report the first stage in table OA.2. The rank condition is satisfied: we find that local migration rates affect the individual probability of moving. Since migration occurs in waves, we find the intuitive result that the likelihood to move is higher if past migration out of the region has been low or future mobility is higher. In terms of the exclusion restriction, one concern might be that outward migration could affect the equilibrium in the local labor markets of origin or destination. However, with almost 100 provinces (and 600 LLS in the ANIA analysis) the contribution of the migration flow from a given province to any other is likely to be negligible.[21]

[20] Mobility out of the region is available starting only in 1995. To compute mobility out of the region before 1995, we take the average of the ratio between the number of movers out of the region and out of the municipality for the overlapping years. This gives the average share of movers out of the region on total movers. We then multiply mobility out of the municipality by this ratio to obtain an estimate of mobility out of the region for the years before 1995.

[21] To assuage further concerns on the validity of the exclusion restriction, we consider two additional exercises. The first is to notice that the model is technically identified even without an exclusion restriction, thanks to the nonlinearity of the inverse Mills ratio. If we drop the exclusion restriction altogether (as a way of addressing an extremely skeptical view of the instrument validity), the effect of entrepreneurial density at learning age is slightly smaller and (as expected) noisier. However, it remains in a similar ballpark and is still significant at the 10% level. A different objection is that there might be correlated shocks for clusters of provinces or regions, which invalidates the identification in the movers model for those who move within the cluster. To address this issue, we restricted both the SHIW sample

The results of the second stage are reported in column 4 of table 2. The coefficient on the Mills ratio is negative and significant, pointing to a negative correlation between the propensity to move and that of becoming an entrepreneur. This is consistent with previous evidence documenting that entrepreneurs are less mobile than employees (Michelacci and Silva 2007). The basic result is qualitatively unchanged, but the coefficient becomes substantially larger—a pattern that will emerge in all of the regressions focusing on movers. This suggests that the movers regression corrects a potential downward bias in the estimates obtained using all workers. A plausible explanation is place-based policies, such as government programs to foster entrepreneurship in places where there is little entrepreneurial activity (i.e., where ED at age 18 is low). These policies may induce a negative correlation between ED at learning age and the decision to become an entrepreneur for stayers, who can benefit from the policy and are thus more likely to become entrepreneurs. Because place-based policies in the learning location apply only to stayers, the bias is absent in the sample of movers.

While our measure of density carries considerable explanatory power, there could be other dimensions of heterogeneity in entrepreneurial density across geographical units that are important for learning, such as the share of patenting firms or the distribution of R&D spending. Unfortunately, data limitations do not allow us to account for these dimensions. However, we can control for average firm size at learning, which tends to be correlated with R&D spending and innovation as well as with market power. If we do so, all our results are unchanged.[22]

Table A1 replicates the same regressions using the more stringent definition of entrepreneurship, which excludes self-employment. This reduces the incidence of entrepreneurship from 19% to 9% (see table 1). Results are qualitatively confirmed. The point estimates are reduced by half (with the exception of the movers sample), which is expected, as the share of entrepreneurs is substantially smaller. According to the estimates of column 2, increasing density at learning age by 1 standard deviation increases the probability that an individual becomes an entrepreneur by 0.8 percentage points, which represents a slightly larger increase relative to the sample mean than in the case of definition 1 (10% vs. 8%).

---

and the ANIA sample to long-distance migrants, i.e., to those who migrate outside the region or outside the macro area (a collection of regions) of birth. We have repeated all of the regressions in the paper involving movers adding these restrictions. The results, reported in table OA.3, fully confirm those of the unrestricted movers sample.

[22] This is true not only for the results in the SHIW sample but also for all those using the ANIA sample discussed in the next sections. The detailed regression results are shown in table OA.4.

B. ANIA Sample: Sectoral Choice

The ANIA sample contains only entrepreneurs and cannot be used to study the occupational choice. However, differently from SHIW, the ANIA data set classifies businesses using a four-digit sector classification. This allows us to study the sectoral choice of activity, conditional on being an entrepreneur.

To study the effect of sectoral density at learning age on the choice of sector in which to operate (conditional on having already chosen to be an entrepreneur), we define sectoral density $\mathrm{ED}SET_{jht} = firms_{jht}/pop_{jt}$ as the ratio of firms in sector h in location j in year t to location j's population in year t. We investigate whether, conditional on becoming an entrepreneur, the probability that an individual operates in sector h increases with sectoral density at learning age. From an econometric point of view, this choice is identical to that of a consumer choosing between H mutually exclusive discrete goods, with sectoral density interpreted as (the inverse of) the "price" of the "good." The idea is that for an entrepreneur located in Silicon Valley the cost of starting an ICT business is lower than the cost of starting a textile business, or the benefits larger. To capture the choice of sector, we use a modified version of equation (5):

$$Z^*_{iht} = \theta + \delta \mathrm{ED}SET_{jh(i,t_L)t_L} + \phi \mathrm{ED}SET_{jh(i,t)t} + D_h + \eta_{iht}, \tag{7}$$

where Z^*_{iht} is the latent net utility associated with running a business in sector h in year t, $\mathrm{ED}SET_{jh(i,t_L)t_L}$ is sectoral density in sector h and area j in which the individual was living at learning age, $\mathrm{ED}SET_{jh(i,t)t}$ is sectoral density in the place where the business is currently located, and D_h are sector dummies. We define indicators $Z_{iht} = 1\{Z^*_{iht} = \mathrm{argmax}\{Z^*_{i1t}, \dots, Z^*_{iHt}\}\}$ for $h = 1, \dots, H$. Assuming that η_{iht} is a type I extreme-value random variable, this type of problem can be analyzed with the conditional logit model of McFadden (1974). Given that all regressions include sector dummies, the parameters are identified only by within-sector, across-year location variation in sectoral density (i.e., identification comes from changes in entrepreneurial density between the year an individual was 18 and the current year of observation). Finally, since ANIA is a cross section of entrepreneurs, $t = 2007$ for all individuals and $\mathrm{ED}SET_{jh(i,t)t}$ varies only by location.

Table 3 reports the results. We find that sectoral density at learning age strongly affects sectoral choice as an entrepreneur (col. 1). The size of the effect decreases somewhat but remains positive, large, and significant when including current sectoral density (col. 2), which, mechanically, is positively associated with individual sectoral choice. To give a sense of the magnitude of the effect, using the estimate in column 2 we find that an increase of sectoral density at learning age of 0.0035, equal to the standard deviation of its distribution, increases the probability that an entrepreneur chooses that sector by 10%.

Table 3
Conditional Logit Model: Probability of Becoming an Entrepreneur in a Given Sector

	(1)	(2)	(3)	(4)	(5)	(6)
EDSET_h learn	46.920***	29.136**	27.948***	20.298*	34.103*	29.244
	(12.204)	(13.345)	10.087)	(12.222)	(19.797)	(19.964)
EDSET_h today		48.990***	30.711***	47.642***		60.295***
		(12.962)	(10.759)	(12.956)		(18.3449)
Number of observations	34,023	34,023	9,279	9,279	2,340	2,340
Number of cases	1,031	1,031	1,031	1,031	260	260
Number of sectors	33	33	9	9	9	9
Individual controls	No	No	No	Yes	No	No
Sample	All	All	All	All	Movers	Movers

NOTE.—Shown are results of a conditional logit regression for the entrepreneurs' sectoral choice. Each observation is an entrepreneur-sector combination. EDSET_h learn is the ratio of the number of firms active in sector h and the local population (sectoral density) at 18 in the place of learning. EDSET_h today is sectoral density in 2007 in the place of residence. In cols. 1–3 the number of sectors is 33, and in cols. 4–6 it is 9. Column 4 includes controls for sex, age, experience, father entrepreneur, years of education, and area of learning as well as area of residence dummies. All regressions include sector dummies.

* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

We next sharpen the specification, first including individual characteristics and then focusing on movers. Since the model with individual characteristics is much more demanding in terms of estimation and the sample includes only 260 movers, we aggregate our original 33 sectors into nine, based on sectoral similarity and sample size considerations. We detail the sectoral aggregation procedure in the appendix.[23] In column 3 we first reestimate the model with the coarser sectoral aggregation on the whole sample. The estimates are similar to those in column 1 with all 33 sectors, with the coefficient smaller for current density (30.7 vs. 48.9). This is to be expected: with more aggregated sectors, the sectoral choice is defined less precisely. In column 4 we estimate a mixed logit model; that is, we add individual controls (sex, age, experience, father entrepreneur, years of education, area dummies). Despite the more demanding specification, results are similar to those in column 2, with the density at learning age dropping from 29 to 20 but maintaining statistical significance at 10%.

As discussed in section III, a problem with the interpretation of these regression coefficients is the usual endogeneity issue. It might be that individuals who chose to start up a textile firm in Prato did so because the textile industry in Prato was particularly profitable at an individual's learning age due to some local characteristics whose effects persist today. As before, we can address this problem by focusing on the sample of movers. In column 5 we run the model on the sample of movers only. The estimate is similar (34.1) and significant at 10%. This confirms that our entrepreneur who grew up in Prato is more likely to run a textile business even if she moves to, say, Parma. In column 6 we also include sectoral density today. As before, the coefficient of sectoral density at learning age drops slightly to 29.2 and, as a consequence, the p-value drops as well (to 14%). Still, the result lines up with the previous ones and confirms that sectoral density at learning age is a strong predictor of entrepreneurs' sectoral choice.

Overall, both the occupational choice results and the sectoral choice results indicate that density at learning age has a strong impact on individual choices. We now move on to analyze the effects of entrepreneurial density on the performance of entrepreneurs.

[23] Our aggregates are (1) agriculture, extraction, and traditional manufacturing (food, tobacco, textile, leather); (2) paper, refinery, and chemicals; (3) metals and nonelectrical machines; (4) electrical and electronic machines, precision instruments, transportation equipment, and other manufacturing NEC (not elsewhere classified); (5) construction; (6) retail trade; (7) hotels and restaurants; (8) transport, communications, and utilities; and (9) other services.

VI. Performance Analysis

As with occupational choice, we exploit both data sets to analyze entrepreneurial performance, starting with the SHIW sample and then moving to the ANIA sample.

A. SHIW Sample: Entrepreneurial Profits

We measure entrepreneurial success with income from business. Since this is available only for entrepreneurs, we estimate a Heckman selection model to correct for selection into entrepreneurship and use as an exclusion restriction the number of family earners, which we assume affects the decision to become an entrepreneur (if other family earners offer some earnings risk diversification) but not entrepreneurial success directly. Otherwise, the set of controls is the same as in the probit estimates in tables 2 and A1. Results are shown in table 4 for the broad definition of entrepreneur. *ED* at learning age has a positive and strongly significant effect on the (log of) entrepreneur's earnings. The effect is slightly smaller but retains fully its statistical significance even after controlling for the current firms density in the province (col. 2). The effect of current *ED* is positive and significant, consistent with a large literature on agglomeration economies (Rosenthal and Strange 2004; Moretti 2011), but the effect of *ED* at learning age is almost twice as large (5.5 vs. 3.0) and is highly statistically significant. In terms of magnitude, the estimate implies that increasing density at learning age by 1 standard deviation increases entrepreneurial income by 11%.

How to interpret these estimates? First, the fact that our proxy for learning opportunities has a positive effect on entrepreneurial performance suggests that "learning skills" (which increases average performance) is more relevant than "learning how to reduce business startup costs" (which lowers it), based on the implications of the simple model of section II. Second, the fact that external effects related to firm density at learning age appear to be more important for firm profits than current externalities is an important result for the literature on agglomeration economies and the channels through which they operate. Duranton and Puga (2004) propose three mechanisms through which agglomeration economies can affect firm performance: the opportunities to learn from other firms, the size of the local workforce (which can increase the division of labor and the quality of job-worker matches), and a greater variety of intermediate inputs. Of these three mechanisms, only learning can have effects that persist once an entrepreneur moves from a high-density area. Our results therefore indicate that learning externalities are at play in the determination of agglomeration economies.[24]

[24] Of course, this does not imply that the other sources of externalities are absent. In fact, *ED* is a natural indicator of learning externalities but not necessarily of the size of the local workforce or of intermediate input varieties, typically captured by other indicators (Glaeser et al. 1992; Cingano and Schivardi 2004).

Table 4
Entrepreneurial Income, Survey of Household Income and Wealth (SHIW) Sample, Definition 1

	(1)	(2)	(3)	(4)
ED learn	6.633***	5.453***	5.404***	6.879**
	(1.138)	(1.310)	(1.474)	(2.976)
ED today		2.974**	3.756**	.790
		(1.273)	(1.670)	(2.820)
LC learn	−.128***	−.114***	−.131***	−.088
	(.030)	(.033)	(.044)	(.064)
LC today		−.069	−.070	.014
		(.063)	(.070)	(.121)
Male	.432***	.433***	.423***	.423***
	(.023)	(.023)	(.023)	(.054)
Age	.271***	.226**	.204**	.385*
	(.091)	(.095)	(.102)	(.208)
Experience	.060**	.058*	.062**	.010
	(.030)	(.030)	(.030)	(.063)
Parent entrepreneur	.091***	.091***	.091***	.121**
	(.025)	(.025)	(.024)	(.059)
Married	.117***	.117***	.113***	.067
	(.026)	(.026)	(.026)	(.059)
Family size	.009	.010	.011	.040**
	(.009)	(.009)	(.009)	(.017)
GDP growth at learning age	.096	.084	.077	−.343
	(.239)	(.238)	(.235)	(.618)
λ_{entr}	.191	.188	.175	.216
	(.028)	(.028)	(.027)	(.139)
λ_{mover}				−.058
				(.079)
Observations	11,408	11,408	11,408	2,009
Area dummies:				
Macro area of birth	X	X		X
Macro area of residence	X	X		X
Region of birth			X	
Region of residence			X	

NOTE.—The dependent variable is log income from entrepreneurial activity. Entrepreneur definition 1 includes (*a*) individual entrepreneurs, (*b*) self-employed/craft workers, (*c*) owner or member of family business, and (*d*) working shareholder/partner. *ED* learn is entrepreneurial density at 18 in the place of birth, and *ED* today is current entrepreneurial density in the place of residence. LC learn and LC today are indicators for liquidity constraints at learning age and in survey year, respectively. All regressions include year, education, and sector dummies. All regressions are the second stage of a Heckman two-stage model to correct for the choice of becoming an entrepreneur. The excluded variable is the number of income recipients in the family. Column 4 uses only the sample of movers, correcting for selection with a Heckman model in which the excluded variables are the average mobility rate out of the region of birth in the 10 years before and after the learning age. Standard errors clustered at the level of year of learning and province are in parenthesis. GDP = gross domestic product.

* Significant at 10%.

** Significant at 5%.

*** Significant at 1%.

Moreover, they also point to a specific mechanism of learning externalities—that is, those that get embedded in the individual's human capital in the coming-of-age years, which are distinct from contemporaneous knowledge spillovers that result from being located in a certain area in the current period.

As for the effect of other controls, we find that male entrepreneurs earn a substantial premium (more than 40%) over female entrepreneurs. Older entrepreneurs earn more, as do those who are married and those with a parent who already was an entrepreneur.[25]

In column 3 we include regional dummies (for both current and learning age location) and find no differences in the estimates, in line with the hypothesis that the *ED* at learning age is not proxying for other unobservables that determine both *ED* at learning age and entrepreneurial success today.

In the last column we focus on the sample of movers, where identification is robust to the concerns discussed in section III. Column 4 corrects for joint selection into entrepreneurship as well as moving, where the instruments for mobility are the same as in the occupational choice regressions of the previous tables. We find that the basic conclusions are confirmed, with the effect becoming slightly larger compared with those for the whole sample.[26]

Table A2 uses the stricter definition of entrepreneurship. This reduces the observations from 11,408 to 5,034. The reduction in sample size affects the statistical power so that, in the specification with regional dummies, we lose statistical significance. However, the pattern that emerges is fully consistent with that based on the broader definition of entrepreneurship. We also find that in the movers sample the coefficient is substantially larger than that in the overall sample, arguably because breaking the correlation between *ED* at learning age and current *ED* is more important in the smaller sample of strictly defined entrepreneurs.

[25] Growing up in an entrepreneurial family can be an important alternative source of learning. Of course, this cannot be told apart from genetic effects, so that one cannot interpret this coefficient as evidence for learning.

[26] It can be argued that using the number of current income earners as an exclusion restriction for entrepreneurship may be invalid if entrepreneurs help their relatives in finding a job more effectively than nonentrepreneurs, e.g., because they can hire them or can facilitate their hiring at other firms. To address this issue we use, as an alternative exclusion restriction, the number of household members aged ≥ 18, ≥ 25, and ≥ 35 at the time the individual started his or her current job (which for the entrepreneur coincides with the year in which the business started). This exclusion restriction captures the same underlying idea (adult members can potentially provide startup capital), but it does not suffer from reverse causality because the firm was not in operation yet. While the resulting sample is much reduced (since the variable that measures when the current job started is missing for half of our sample), the results remain reassuringly similar. The coefficient on the firm density at learning age is 6.03 (significant at 10%), very similar to the 6.88 estimate in col. 4 of table 4.

B. ANIA Sample: Productivity Performance

To further investigate the effects of *ED* on entrepreneurial quality, we now turn to the ANIA sample. We use the same framework as in the SHIW sample, with a few exceptions. First, following Lucas (1978), we measure entrepreneurial ability with firm TFP. Second, we measure *ED* in the location in which the entrepreneur was actually living at age 18 rather than at birth. As stated above, for 15% of cases in our sample these two locations differ. Third, the set of controls is the same as in the SHIW regressions, with the exception that in ANIA we do not have total experience and hence use the number of years since the individual started managing the firm. As before we cluster standard errors at the learning age–LSS level, and as before results are robust to alternative clustering schemes.

Table 5 shows the results. *ED* at learning age has a positive and precisely estimated effect on the firm TFP (col. 1). Increasing *ED* at learning age by 1 standard deviation (0.02) increases TFP by 8.6%. None of the other controls displays a significant effect, likely because we have a much smaller sample compared with SHIW. It is hence remarkable that *ED* at learning age is also statistically significant in this data set. In column 2 we add current *ED*, for which we find a positive but statistically insignificant coefficient. The coefficient on *ED* at learning age increases slightly (to about 5) and remains statistically significant at conventional levels. As in the SHIW data, therefore, we find the striking result that learning age externalities play a stronger role than current production externalities. We also notice that the elasticity of entrepreneurial quality to *ED* is very close in the two data sets, although the sampling frame and the variables used to measure entrepreneurial performance are different. Once more, the results suggest that the skill improvement effect dominates the entry cost reduction effect mentioned in section II.[27]

The remaining columns of table 5 perform a series of robustness checks. First, we control for *ED* at birth (col. 3). A growing literature stresses the role of early education on professional outcomes (Heckman, Pinto, and Savelyev 2013). Extrapolating from this literature, it may be argued that the economic environment that matters most to accumulate entrepreneurial ability is the one in the early years of development. However, we find that the estimated

[27] The ANIA sample allows us to test robustness to the issue of parental sorting discussed in sec. III. Parental sorting could generate a spurious correlation between performance and firm density at learning age if parents move to locations with better learning opportunities to indulge their children's entrepreneurial attitudes. One simple test for this is to reestimate the model on a sample where parental sorting is absent, i.e., to exclude those who moved between birth and age 18 (since in this age interval children move because their parents move). If we focus on this restricted sample, we find that the effect of firm density on firm performance is 4.3, only slightly smaller than when the full sample is considered (table 5, col. 2). This suggests that parental sorting is unlikely to drive our results.

Table 5
Total Factor Productivity (TFP) and *ED* at Learning Age, Associazione Nazionale delle Imprese Assicurative (ANIA) Sample

	(1)	(2)	(3)	(4)	(5)
ED learn	3.680*	4.711**	5.494**	5.794**	11.378***
	(1.922)	(2.239)	(2.373)	(2.873)	(3.653)
ED today		2.779	2.904	1.945	2.684
		(2.584)	(2.629)	(4.551)	(4.586)
LC learn	−.071	.018	.017	.142	.167
	(.103)	(.108)	(.109)	(.249)	(.170)
LC today		−.420**	−.416**	−1.245	−.421
		(.204)	(.207)	(.842)	(.356)
Male	−.065	−.072	−.070	−.041	−.037
	(.048)	(.048)	(.049)	(.052)	(.098)
Age	.298	.425*	.369	.483	.384
	(.239)	(.243)	(.250)	(.408)	(.486)
Experience	.033	.037	.042	.041	−.041
	(.033)	(.032)	(.033)	(.034)	(.058)
Parent entrepreneur	−.068	−.072	−.075	−.082	.155
	(.051)	(.052)	(.052)	(.055)	(.110)
Married	.094	.094	.100	.098	.030
	(.068)	(.067)	(.068)	(.072)	(.138)
Family size	−.007	−.007	−.011	−.008	.000
	(.023)	(.023)	(.023)	(.024)	(.045)
GDP growth at learning age	−2.445	−2.500	−2.501	−1.963	9.606
	(3.298)	(3.306)	(3.323)	(3.989)	(6.145)
ED birth			−3.041		
			(3.253)		
λ_{mover}					.142
					(.233)
Observations	2,531	2,531	2,495	2,531	637
Area dummies:					
Macro area of residence	X	X	X		X
Macro area of learning	X	X	X		X
Province of residence				X	
Province of learning				X	

NOTE.—The dependent variable is log TFP. All regressions are based on the data for the years 2005–7 and include year, education, and sector dummies. Column 5 uses only the sample of movers, correcting for selection with a Heckman model in which the excluded variable is the average mobility rate out of the municipality where the individual was living at learning age in the 10 years after the learning age. Standard errors clustered at the level of the local labor system and year of learning are in parenthesis. GDP = gross domestic product.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

coefficient of *ED* at birth is not statistically different from zero while that of *ED* at learning age is unaffected, supporting the idea that most of the "learning from other entrepreneurs" occurs after childhood.

These findings could still be consistent with a cultural explanation if the culture that matters for entrepreneurship is not acquired in the early years of life but later. If the density where one grows up is a reflection of an underlying entrepreneurial culture, our measure might just be proxying for the latter rather than for learning opportunities. As before, we address this concern relying on the idea that culture evolves slowly and the geographical unit that is covered by culture is broader than that where learning entrepreneurial abilities from firms takes place. Accordingly, in column 4 we expand the number of spatial controls to account for potential spatially correlated effects. Given that the geographical unit for *ED* is the LLS, we can use finer controls than those of the SHIW sample and insert 95 province dummies (rather than 20 regional dummies). Since provincial governments are in charge of managing schools, provincial dummies also account for any persistent geographical differences in the quality of schooling. As before, we use separate dummies for location at learning age and current location. Again, we find that the point estimates are unchanged—if anything, they become larger—and that the coefficient is statistically significant at the 5% level.[28]

Finally, we focus on the sample of movers employing a Heckman sample selection model. Compared with the SHIW regressions, we introduce two slight modifications given that we have richer data. First, we use the rate of mobility out of the municipality rather than out of the region. In fact, an LLS contains on average 10 municipalities, while regions contain on average around 40 LLS. Mobility out of the municipality is therefore a closer proxy to mobility out of the LLS than mobility out of the region. Second, we only use the average mobility rate in the 10 years after learning age. Different from the SHIW sample (where we observe the province of birth but not the province of residence at learning age), in the ANIA data we know that the individual was still resident in the LLS of learning at 18, so we can disregard prior mobility. The first stage, reported in table OA.2, shows that the instrument has the expected positive sign and is statistically significant. Column 5 shows that, as in the SHIW sample, the effect of *ED* at learning age becomes stronger and more precisely estimated. In this case, the Mills ratio is not significant; consistent with this, a regression without the Heckman correction delivers very similar results.

We refine the test further by including a full set of current province-time fixed effects and even the fully saturated current LLS-time fixed effects. This way, we control for any current time-varying local shock possibly correlated with current *ED*, which, as noted in footnote 18, could transmit the bias to *ED* at learning age.[29] This refinement is particularly powerful for

[28] We have also experimented with dummies at the LLS level, thus exploiting for identification of learning effects only on the time variation in *ED*. As with the SHIW data, estimates for the effect of *ED* at learning age are similar in magnitude but with larger standard errors. This is not surprising, given the large number of dummies (more than 400) and the limited sample size.

[29] Unfortunately we could not implement this strategy in the SHIW sample because of convergence problems when estimating probit models with more than 1,000 additional dummies (11 waves times 95 provinces).

movers, where we compare individuals who are today in the same province (LLS) but grew up in different provinces (LLS) and, in light of the controls, make the same decision with regard to where to move to. We find that results for the basic specifications in table 5 are if anything strengthened, regardless of whether we use province-time or LLS-time fixed effects. They are confirmed even in the case in which we allow the dummies to differ for young and old entrepreneurs, to control for some secular trends in the local effects. We report the results in table OA.5.

As a final check, we have used sales per worker to measure entrepreneurial quality. While TFP is the closest empirical counterpart to ability in the Lucas model, estimating TFP requires more assumptions than just computing a simple measure of labor productivity. Results in table A3 are similar to those obtained using TFP. The coefficient on *ED* at learning age tends to be larger and more precisely estimated. Increasing density at learning age by 1 standard deviation results in an increase in sales per worker of around 22%. A possible explanation is that growing up in denser areas not only improves ability but also leads entrepreneurs to increase the capital-labor and intermediate inputs–labor ratio, leading to an even stronger effect on labor productivity.

Thus far, we have taken age 18 as a focal point for learning age. This is partly because of data limitations (in ANIA we observe the place of residence at birth, at age 18, and at the time of the interview) and partly because age 18 should capture learning occurring before an individual makes her formal occupational choice. However, it is likely that—as argued in section II.B—learning occurs over an age interval, including periods in which the individual works for firms different from the one she ultimately creates. To shed some light on this process, we reestimate the relationship between firm TFP and density at learning age in the ANIA data (using the basic specification in col. 2 of table 5) assuming that the learning age ranges between 10 and 25 (and that individuals who are in location j at 18 were in that location throughout this particular segment of their life cycle).[30] Interestingly, the effect is small and insignificant at age 10, and it increases monotonically with age until it reaches a peak at age 23 (see fig. A2). Given that most of the entrepreneurs in our sample were already working at age 20 (i.e., after completing high school), the fact that the coefficient keeps increasing up to age 23 is consistent with the idea that learning also occurs in firms.

Overall, we take this evidence as supportive of the idea that one important channel through which individuals acquire entrepreneurial abilities is early exposure to a richer entrepreneurial environment.[31]

[30] We do this exercise only on the ANIA data, since for SHIW we know only location of residence at birth.

[31] In light of evidence arguing that low- and high-skilled entrepreneurs create firms of different types (Gendron-Carrier 2018), we explore whether the effect of ED at learning age varies by level of schooling. In both SHIW and ANIA, we

VII. What Features of Entrepreneurship Are Learnable?

Being an entrepreneur requires multiple talents. The entrepreneur develops new ideas, evaluates their market appeal, organizes production, bears the risk of failure, and so forth. The role of the entrepreneur as the bearer of risk dates back to Knight (1921), who ascribes the very existence of the firm to its role as an insurance provider. More generally, there is a large managerial literature showing that entrepreneurs tend to differ from the average individual in terms of some key traits, such as confidence and optimism (Camerer and Lovallo 1999; Åstebro 2003). Besides possessing these traits, the entrepreneur also needs to be able to conceive and implement ideas, arrange and coordinate the production process, and bring the product to market. Marshall (1890) was the first to stress the importance of localized spillovers to learn "the mysteries of the trade."

In this section we offer empirical evidence on how learnable these features are, using information from the ANIA survey. This is a simple way of looking at the "black box" of what our variable of interest (entrepreneurial density at learning age) represents. In face-to-face interviews, respondents were asked a set of questions aimed at eliciting risk preferences and identifying personality traits. We briefly describe them here and report the full questions in the appendix. First, respondents were asked to choose between different investment strategies with decreasing risk-return profiles, ranked from 1 (high risk and high return) to 5 (low risk and low return). We use the answers to this question as a measure of risk aversion. Second, the respondents were asked to express, on a scale from 1 to 5, their preferences regarding drawing a ball from an urn with 50 green and 50 yellow balls versus drawing it from an urn containing an unknown share of balls of each color. We use this question to measure ambiguity aversion. A measure of self-confidence was obtained by asking the entrepreneur whether she ranked herself below, at the same level, or above the average ability of other entrepreneurs. Optimism is measured by the answer (on a scale from 0 to 10) to how much the respondent agrees with the statement "All things considered I

find that the effect of firm density at learning age on performance declines with education, consistent with college graduates learning less. On the other hand, the interaction with education has the opposite effect on the probability of becoming an entrepreneur: an increase in firm density at learning age increases the probability of becoming an entrepreneur more for the college educated (table OA.6). Density may also have heterogeneous effects across sectors. Grouping observations in two broad sectors—industry (manufacturing, construction, and utilities) and services (so that we have enough observations in each)—we test whether density matters differently in these two sectors. The results are not clear-cut. In the SHIW sample we find some evidence that firm density at learning age has a greater impact for occupational choice in services (though this difference disappears in the sample of movers). In contrast, in both the SHIW sample and the ANIA sample we find no statistically significant evidence that density affects performance deferentially in the two sectors (see table OA.7).

Table 6
Descriptive Statistics: Traits and Managerial Practices

Variable	Obs	Mean	SD	Min	Max
	A. Traits				
Risk aversion	967	2.70	.67	1	4
Ambiguity aversion	955	3.40	1.44	1	5
Confidence	944	2.18	.40	1	3
Optimism	946	7.33	1.76	0	10
Satisfaction	937	7.47	1.67	0	10
	B. Managerial Practices				
Management	388	2.41	.66	1	4.41
Monitoring	388	2.67	.82	1	5
Targets	388	2.41	.90	1	5
People	388	2.33	.64	1	5

NOTE.—See sec. A4 of the appendix for definitions of variables.

expect more good than bad things in life." Job satisfaction was measured from the answer to the question "Excluding the monetary aspects and considering only the other characteristics of your job, can you tell me if they give you more satisfaction or annoyance?" again on a scale of 0 to 10. Summary statistics for these variables are reported in panel A of table 6.

The original survey lacks information on the ability to coordinate production or manage the firm. In the fall of 2012 we recontacted the entrepreneurs in the ANIA sample and asked them to participate in a second round of interviews to assess their managerial practices. We used the methodology developed by the World Management Survey (WMS; see Bloom and Van Reenen 2010a, 2010b). The WMS is based on a telephone double-blind survey technique and comprises a set of open-ended questions whose qualitative answers are then recoded into quantitative measures with a score ranging between 1 (worst managerial practices) to 5 (best managerial practices). The questionnaire comprises five sections that consider different key areas of management practices. We investigate three areas. The first section is "Monitoring" and focuses on the monitoring of performance and reviewing the results. The second section is "Targets" and aims at assessing the respondents' managerial ability in identifying quantitative and qualitative targets, their interconnection, and their temporal cascade. The third section is "People," and it is concerned with human resource management, ranging from promoting and rewarding employees on the basis of performance, removing poor performers, and hiring and retaining the best workers. The average score of the three areas defines the index of overall managerial ability.[32] Of the original

[32] The original WMS contains two additional areas, operations and leadership. These areas investigate practices that are very sector specific, such as the operation of the production unit. Given that we have firms from both manufacturing

966 entrepreneurs, we were able to reinterview 388 (details on the data collection methodology are provided in sec. A3 of the appendix). Descriptive statistics are provided in panel B of table 6. Given that for this analysis the number of observations drops substantially and given that what we are trying to capture here is more elusive than profits or productivity, we adopt a more parsimonious regression specification, dropping variables that are less important for these outcomes and not significant when included (namely, the married dummy, family size, and real GDP growth at learning), keeping only experience (and dropping age given high collinearity), and aggregating education from five to three categories (elementary and junior high school, high school, and college and postgraduate). Finally, given that in the previous evidence we have found that skills accumulation depends on overall density, in what follows we focus on this variable and disregard sectoral density.[33]

As a first check, in table A4 we regress TFP on personality traits. We have no priors on how the degree of risk aversion or optimism should correlate with firm-level TFP. Indeed, the correlation is always insignificant. Next, we check whether these traits are correlated with *ED* at learning age. The results are reported in table 7. In general, we find no evidence that entrepreneurial traits are affected by *ED* at learning age. Growing up in areas with greater firm density seems to lead to higher risk aversion (col. 1), but precision is low and the estimate is not statistically significant. This evidence is consistent with the predominant role of the innate component in explaining individual risk preferences found by Cesarini et al. (2009) using twin studies. All other traits also appear to be uncorrelated with *ED* at learning age (cols. 2–5). Overall, these results suggests that traits are not affected by the entrepreneurial context at learning age, arguably because these traits are mostly genetically determined or acquired much earlier in life, including at school (Heckman, Pinto, and Savelyev 2013).

The ability to manage a business, however, is more likely to be learnable. Indeed, managerial skills are precisely what one learns in business schools and colleges specialized in teaching entrepreneurship. But these skills may possibly be learned before college by direct observation and exposure to adopted practices by the firms in the place where one grows up.

We test this potential learning channel using the measures of managerial practices discussed above. To first validate these measures, in table A5 we report the results of regressions in which productivity and firm size (measured

and services, we only investigated areas where practices are sufficiently similar to allow us use of the same interview scheme. We control for sector dummies in all regressions.

[33] In unreported regressions we have also investigated the effects of sectoral density, always finding that it has no impact on any of the variables analyzed in this section.

Table 7
Entrepreneurial Traits and *ED* at Learning Age, Associazione Nazionale delle Imprese Assicurative (ANIA) Sample

	Risk Aversion (1)	Ambiguity Aversion (2)	Confidence (3)	Optimism (4)	Satisfaction (5)
ED learn	.277	−3.404	−.991	2.978	2.743
	(1.489)	(3.184)	(.885)	(3.935)	(3.721)
ED today	.750	−4.426	−1.232	−5.968	−6.439
	(2.031)	(4.332)	(1.205)	(5.326)	(5.055)
Male	−.045	.146	.106***	−.056	−.088
	(.048)	(.103)	(.029)	(.127)	(.120)
Experience	.056**	.020	−.014	.107	.054
	(.028)	(.061)	(.017)	(.075)	(.071)
Parent entrepreneur	−.038	−.049	.045	−.012	.089
	(.046)	(.099)	(.028)	(.122)	(.116)
Junior/high school	.129	.087	−.013	−.395*	−.114
	(.086)	(.183)	(.051)	(.225)	(.215)
College/postgraduate	−.032	.081	.099***	.079	.007
	(.053)	(.112)	(.031)	(.138)	(.131)
Observations	967	955	944	946	937

NOTE.—The dependent variables are listed in the first row. All regressions include macro area of learning and of current location as well as sector dummies. Standard errors clustered at the level of the local labor system and year of learning are in parenthesis.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

by the log of the number of employees) are regressed on the managerial practice scores. As shown by Bloom and Van Reenen (2007) and Bloom, Sadun, and Van Reenen (2012b), managerial practices are strongly correlated with the size of the firm and its productivity. This holds in our sample too: each type of managerial practice, as well as the general index, is positively and significantly correlated with TFP and firm size.

Table 8 runs regressions of these measures on *ED* at learning age, controlling for current density. Interestingly, density at learning age has a positive and statistically significant correlation with all of the measures of managerial practices, while current density has an insignificant effect. To give a sense of the magnitude of the effects, increasing *ED* at learning age by 1 standard deviation (0.019) would increase the overall management score by approximately 0.1, which is equal to 6% of the standard deviation of the management score. The effects for the subcomponents of the management score are similar. Among the other controls, we find that males have higher managerial scores, that experience is negatively correlated with the score, and that highly educated individuals have higher scores.

Overall, we take this evidence as suggesting that managerial capabilities can be learned from other entrepreneurs, while traits might be more innate

Table 8
Managerial Practices and *ED* at Learning Age, Associazione Nazionale delle Imprese Assicurative (ANIA) Sample

	Management (1)	Monitoring (2)	Targets (3)	People (4)
ED learn	5.019**	6.461**	5.242*	4.785**
	(2.307)	(2.956)	(3.137)	(2.278)
ED today	−.422	−1.687	−2.951	1.599
	(3.063)	(3.926)	(4.165)	(3.025)
Male	.269***	.307***	.270***	.223***
	(.075)	(.096)	(.102)	(.074)
Experience	−.108**	−.140**	−.178**	−.035
	(.054)	(.069)	(.074)	(.053)
Parent entrepreneur	−.106	−.078	−.121	−.146**
	(.068)	(.087)	(.093)	(.067)
Junior/high school	−.207	−.122	−.325	−.164
	(.146)	(.187)	(.198)	(.144)
College/postgraduate	.187**	.145	.285***	.215***
	(.076)	(.097)	(.103)	(.075)
Observations	386	386	386	386

NOTE.—The dependent variables are listed in the first row. All regressions include macro area of learning and of current location as well as sector dummies. Standard errors clustered at the level of the local labor system and year of learning are in parenthesis.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

and therefore less subject to the influence of the economic environment: would-be entrepreneurs learn from other entrepreneurs around them how to run a business.

VIII. Conclusions

This paper analyzes the extent to which growing up in a high entrepreneurial geographical area increases both the likelihood that an individual becomes an entrepreneur and her entrepreneurial quality or success. We find evidence that this is indeed the case, as would be implied by models where entrepreneurial ability is socially acquired. Interestingly, we find that the effect of entrepreneurial density at learning age is stronger (quantitatively and statistically) than that of current entrepreneurial density, which captures more traditional spillover effects. A remarkable finding is that the results we find hold in two distinct data sets and are robust to a large set of controls. Moreover, we find evidence that individuals growing up in high firm density areas acquire managerial skills but that individual traits reflecting risk aversion, aversion to ambiguity, and self-confidence or optimism (which are traditionally associated with entrepreneurship) are independent of location. This suggests that the "personality traits" factor of entrepreneurship has a larger innate component, swamping any effect played by the environment.

Appendix

Data Details

A1. Sectoral Concordance for the 1951–2011 Census

The National Institute of Statistics updates the sectoral classification of firms at every census, which is held every 10 years. The sectoral classification has undergone substantial changes in the 60-year period we consider. As a consequence, the finest homogeneous classification we can construct comprises 33 sectors. We now give some further details on the construction of the concordance.

- Data from 1951 to 1991 are supplied in a unique file with a homogeneous classification. The only exception is that "clothing and shoes" are lumped together in 1951. To separate them, we take the respective shares in 1961 and apply them to the aggregate value in 1951. For example, if in a locality in 1951 there were 100 firms in "clothing and shoes" and in 1961 there were a total of 150 with 50 in "shoes," we assume that in 1951 there were 33 shoemakers. The same holds for the machines sector, which is more aggregated in 1951 than in the successive census. We apply the same procedure to disaggregate the 1951 data.
- In 2001 the sectoral classification is Ateco91 at six digits, which can be easily matched to that of 1951–1991.
- In 2011 the sectoral classification is Ateco07 at three digits.
- In the firm data the sectoral classification is Ateco91 at six digits.

The final sectoral list is the following: agriculture, mining, food and beverages, tobacco, textile, apparels, wood and furniture, paper, printing and publishing, chemicals, plastic, nonmetallic minerals, metallic minerals, nonelectric machines, electric machines, precision instruments, transportation equipment (including cars), manufacturing not otherwise classified, utilities, water, garbage collection, construction, retail trade, wholesale trade, hotels and restaurants, transport, post and telecom, finance, insurance, business services, entertainment, repairing, and public administration.

A2. The ANIA Survey

The ANIA Survey for Small Business Companies collects data on a sample of 2,295 Italian firms and their top manager. The survey was conducted on a sample of small Italian firms, having up to a maximum of 250 employees, extracted from the total number of companies registered with the Cerved Group, a business information agency operating in Italy that collects companies' balance sheet data. The survey was conducted between October 2008 and July 2009.

Compared with the initial target set at the completion of 2,300 interviews, the investigation closed with 2,295 completed interviews. Participation in the survey entails the willingness to provide detailed information regarding many aspects of the firm's operations and characteristics as well as the willingness of the chief executive officer (CEO) or owner of the company to take part in a face-to-face interview with a professional interviewer. The first type of data was collected through a questionnaire filled out by each company, while the second type was obtained through an interview using the CAPI method. Partly because the survey took place during the financial crisis and partly because interviews targeted the CEO of the firm, the dropout rate was relatively high, particularly among firms in the larger size categories. To account for this, the survey design was slightly reviewed to include a larger number of smaller firms (with less than 20 employees), which were easier targets. This has caused the sample to be somewhat biased toward smaller firms than the population of businesses with up to 250 employees.

A3. The Managerial Practices Survey

The data collection is based on a telephone double-blind survey technique and comprises a set of open-ended questions that are subsequently evaluated using a scoring grid. Qualitative answers are then recoded into quantitative measures with a score ranging between 1 (worst managerial practices) to 5 (best managerial practices). We first selected a group of interviewers, trained them with a specific program, and then had them run the survey. The data collection process was carried out using the methodology of the World Management Survey (http://www.worldmanagementsurvey.com/). See Bloom and Van Reenen (2010a) and Bloom et al. (2012a) for a full exposition of the survey characteristics and the data collection method.

Not all firms that were recontacted participated in the management survey. A comparison of the observable characteristics of those who refused to participate and those who participated shows no systematic differences in terms of firm characteristics, sector, area of location, and learning. Some small but statistically significant differences emerge in terms of entrepreneurs' characteristics, such as education and having at least one parent who was an entrepreneur (see table OA.8).

A4. Definition of Variables

Here we provide a detailed description of the variables used in the paper.

- *Entrepreneur, definition 1*: This is a broader definition that includes the self-employed, partners of a company, and owners who run an incorporated business.

- *Entrepreneur, definition 2*: This is a narrower definition that includes only partners of a company and owners who run an incorporated business.
- *ED learn*: Entrepreneurial density (number of firms per capita) in the location where the entrepreneur was living at age 18.
- *ED today*: Current entrepreneurial density (number of firms per capita) in the location where the entrepreneur (or the firm in the case of the ANIA sample) is located.
- *EDSET learn*: Entrepreneurial density at the sectoral level (number of firms at the sectoral level per capita) in the location where the entrepreneur was living at age 18.
- *EDSET today*: Current entrepreneurial density at the sectoral level (number of firms at the sectoral level per capita) in the location where the entrepreneur (or the firm in the case of the ANIA sample) is located.
- *LC learn*: Log of bank branches per capita in the location where the entrepreneur was living at age 18. In table 1 we report the mean and standard deviation of the level (rather than the log) of branches per 1,000 inhabitants.
- *LC today*: Log of current bank branches per capita in the location where the entrepreneur (or the firm in the case of the ANIA sample) is located. In table 1 we report the mean and standard deviation of the level (rather than the log) of branches per 1,000 inhabitants.
- *Number of income recipients*: Members of the household who receive some income.
- *Mover*: A dummy equal to 1 if the current location of the individual is different from that where she was born (SHIW sample) or was living at 18 (ANIA sample). In the SHIW sample, we distinguish between all individuals (mover) and entrepreneurs (mover-entrepreneur).
- *Experience*: Labor market experience in the SHIW data and number of years since the entrepreneur has acquired the responsibility of the management of the firm in the ANIA data.
- *GDP growth at learning age*: Regional GDP growth in the region where the entrepreneur was living at age 18.
- *Risk aversion*: Indicator obtained using the answers to the question "If the investment strategy of the firm depends only on you, among the following alternative strategies which one would you pick up? One that yields (*a*) low profits but no risk of losses; (*b*) decent profits and rare losses; (*c*) good profits with some chances of incurring losses; (*d*) very high profits with a high risk of significant losses." The indicator is coded between 1 and 4, increasing in risk aversion.
- *Ambiguity aversion*: Indicator obtained using the answers to the question "Think about two urns, each containing 100 balls, either green or yellow. You win 1,000 euros if you draw an urn of the color of your choice. Choose the color. Urn 1 contains both green and yellow balls,

in unknown proportion. Urn 2 contains 50 green and 50 yellow balls. From which urn would you rather draw the ball? (*a*) Strong preference for urn 1; (*b*) slight preference for urn 1; (*c*) indifferent; (*d*) slight preference for urn 2; (*e*) strong preference for urn 2." The indicator is coded between 1 and 5, increasing in ambiguity aversion.

- *Confidence*: Indicator obtained using the answers to the question "With respect to the average ability of other entrepreneurs, in your job, do you believe to be (*a*) below average; (*b*) average; (*c*) above average." The indicator is coded between 1 and 3, increasing in self-confidence.
- *Optimism*: Answers to the question borrowed from a standard Life Orientation Test (Scheier et al. 1994): "How much do you agree with the statement: Overall I expect more good things than bad things to happen to me." Coded between 0 and 10, with a higher number indicating more optimism.
- *Satisfaction*: Indicator obtained from the following question: "Notwithstanding the profits motive and only considering the other characteristics of your job, can you tell me if they give you more satisfaction or dissatisfaction?" Coded between 0 and 10, with a higher number indicating more satisfaction.
- *Monitoring*: Score variable that assesses the quality of the practices in place within the firm in terms of monitoring of performance and reviewing the results. It takes a value between 1 (minimum quality) and 5 (maximum quality).
- *Targets*: Score variable that assesses the quality of the practices in place within the firm in terms of identifying quantitative and qualitative targets, their interconnection, and their temporal cascade. It takes a value between 1 (minimum quality) and 5 (maximum quality).
- *People*: Score variable that assesses the quality of the practices in place within the firm in terms of human resources management, ranging from promoting and rewarding employees on the basis of performance, removing poor performers, and hiring and retaining the best workers. It takes a value between 1 (minimum quality) and 5 (maximum quality).
- *Management*: Overall score variable obtained as the average of monitoring, targets, and people. It takes a value between 1 (minimum quality) and 5 (maximum quality).

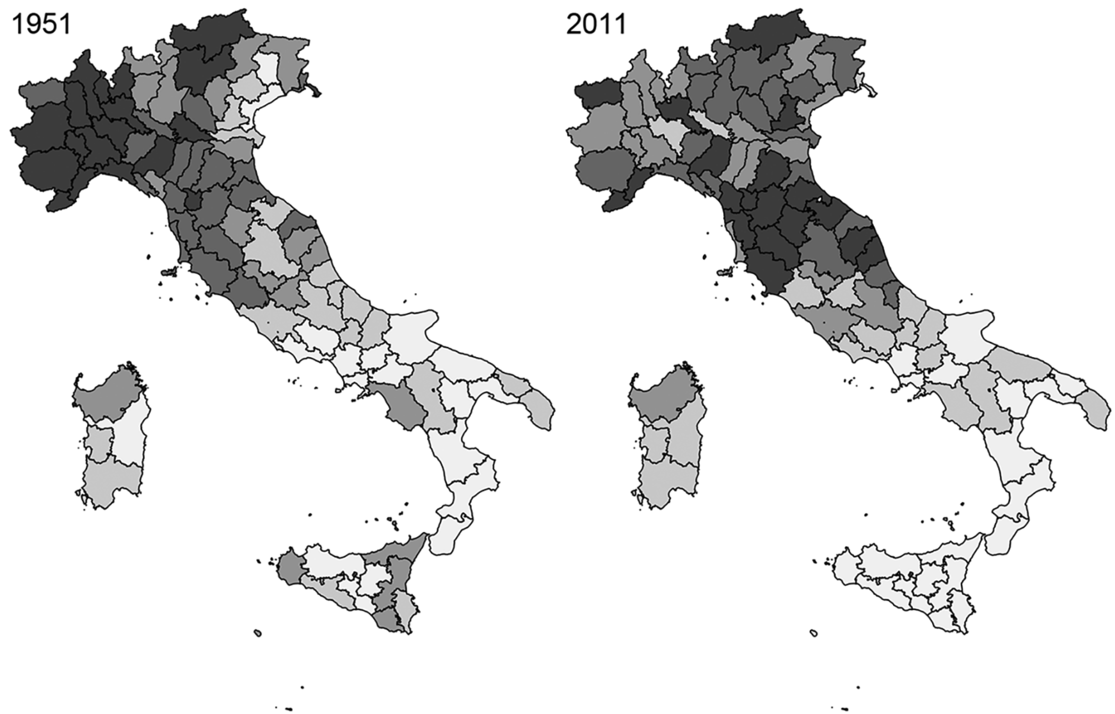

Fig. A1.—Spatial distribution of firm density, 1951 and 2011. The data are taken from the 1951 and 2011 census of businesses. The darkest color corresponds to the top 20%, and the lightest corresponds to the bottom 20%. A color version of this figure is available online.

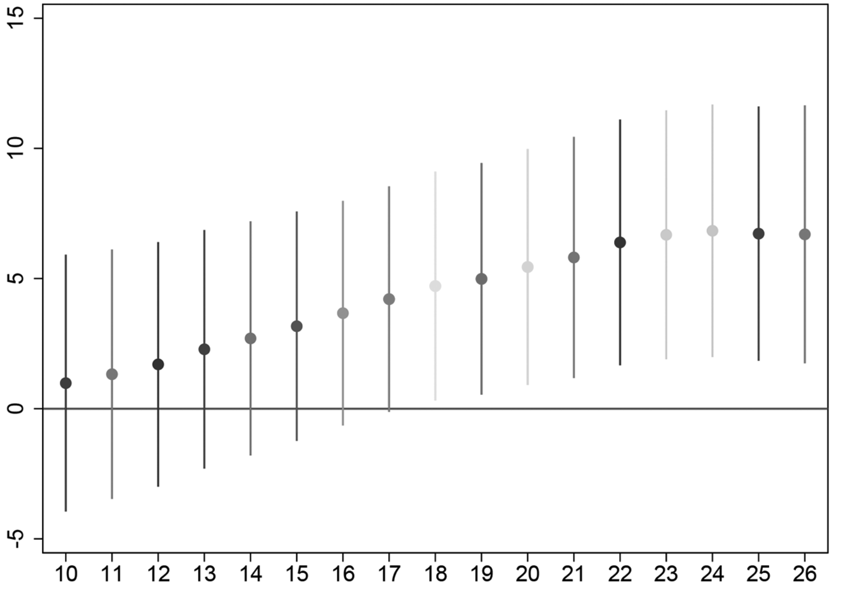

FIG. A2.—Estimates of *ED* learn coefficients at different learning ages. This figure reports the estimated coefficient and 95% confidence interval of entrepreneurial density at ages 10–25, where each estimate is obtained in a separated regression. A color version of this figure is available online.

Table A1
Probability of Becoming an Entrepreneur, Survey of Household Income and Wealth (SHIW) Sample, Definition 2

	(1)	(2)	(3)	(4)
ED learn	.776***	.387*	.283	4.707***
	(.195)	(.222)	(.259)	(.831)
ED today		.821***	1.354***	−1.281
		(.233)	(.290)	(.938)
LC learn	−.008	−.008	−.005	−.018
	(.005)	(.006)	(.008)	(.022)
LC today		−.005	−.008	−.110**
		(.012)	(.013)	(.047)
Male	.009***	.010***	.010***	.026**
	(.003)	(.003)	(.003)	(.012)
Age	.095***	.078***	.073***	.187***
	(.015)	(.016)	(.018)	(.066)
Experience	−.009*	−.010**	−.009*	−.014
	(.005)	(.005)	(.005)	(.015)
Parent entrepreneur	.092***	.091***	.091***	.094***
	(.006)	(.006)	(.006)	(.036)
Married	.026***	.026***	.025***	.011
	(.004)	(.004)	(.004)	(.026)
Family size	.000	.001	.001	.004
	(.002)	(.002)	(.002)	(.005)
GDP growth at learning age	−.067	−.073	−.059	−.134
	(.044)	(.044)	(.044)	(.186)
Number of income recipients	.018***	.018***	.017***	.027**
	(.002)	(.002)	(.002)	(.011)
λ_{mover}				−.882***
				(.015)
Observations	56,292	56,292	56,292	12,244
Area dummies:				
Macro area of birth	X	X		X
Macro area of residence	X	X		X
Region of birth			X	
Region of residence			X	

NOTE.—Shown are results of probit regressions for the choice of being an entrepreneur, marginal effects. Entrepreneur definition 2 includes (*a*) individual entrepreneurs, (*b*) owner or member of family business, and (*c*) working shareholder/partner and excludes (*d*) self-employed/craft workers. *ED* learn is entrepreneurial density at 18 in the place of birth, and *ED* today is current entrepreneurial density in the place of residence. LC learn and LC today are indicators for liquidity constraints at learning age and in survey year, respectively. All regressions include year, education, and sector dummies. Column 4 only uses the sample of movers, correcting for selection with a Heckman model in which the excluded variables are the average mobility rate out of the region of birth in the 10 years before and after the learning age. Standard errors clustered at the year of learning and province level are in parenthesis. GDP = gross domestic product.

* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Table A2
Entrepreneurial Income, Survey of Household Income and Wealth (SHIW) Sample, Definition 2

	(1)	(2)	(3)	(4)
ED learn	5.561***	3.892*	2.592	8.842*
	(1.839)	(2.116)	(2.414)	(4.790)
ED today		4.617**	8.330***	.836
		(2.189)	(2.926)	(4.399)
LC learn	−.092*	−.068	−.071	−.109
	(.049)	(.056)	(.070)	(.116)
LC today		−.121	−.154	.276
		(.100)	(.118)	(.169)
Male	.396***	.396***	.385***	.287***
	(.035)	(.035)	(.034)	(.078)
Age	.505***	.440***	.364**	.882**
	(.153)	(.161)	(.175)	(.360)
Experience	.025	.023	.026	−.009
	(.048)	(.048)	(.049)	(.103)
Parent entrepreneur	.201***	.201***	.202***	.309***
	(.048)	(.048)	(.046)	(.116)
Married	.059	.060	.055	−.017
	(.046)	(.046)	(.046)	(.118)
Family size	.008	.009	.007	.108***
	(.015)	(.015)	(.015)	(.030)
GDP growth at learning age	−.723*	−.730**	−.732**	−1.119
	(.372)	(.369)	(.367)	(1.038)
λ_{entr}	.414***	.410***	.392***	.618***
	(.052)	(.052)	(.051)	(.233)
λ_{mover}				−.002
				(.110)
Observations	5,034	5,034	5,034	906
Area dummies:				
Macro area of birth	X	X		X
Macro area of residence	X	X		X
Region of birth			X	
Region of residence			X	

NOTE.—The dependent variable is log income from entrepreneurial activity. Entrepreneur definition 2 includes (*a*) individual entrepreneurs, (*b*) owner or member of family business, and (*c*) working shareholder/partner and excludes (*d*) self-employed/craft workers. *ED* learn is entrepreneurial density at 18 in the place of birth, and *ED* today is current entrepreneurial density in the place of residence. LC learn and LC today are indicators for liquidity constraints at learning age and in survey year, respectively. All regressions include year, education, and sector dummies. All regressions are the second stage of a Heckman two-stage model to correct for the choice of becoming an entrepreneur. The excluded variable is the number of income recipients in the family. Column 4 uses only the sample of movers, correcting for selection with a Heckman model in which the excluded variables are the average mobility rate out of the region of birth in the 10 years before and after the learning age. Standard errors clustered at the year of learning and province level are in parenthesis. GDP = gross domestic product.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Table A3
Sales per Worker and *ED* at Learning Age, Associazione Nazionale delle Imprese Assicurative (ANIA) Sample

	(1)	(2)	(3)	(4)	(5)
ED learn	11.392***	10.383***	8.742***	9.155**	18.455***
	(2.738)	(3.140)	(3.353)	(4.000)	(5.326)
ED today		.662	.139	4.234	−3.376
		(3.614)	(3.645)	(5.895)	(6.933)
LC learn	−.198	−.230	−.235*	−.547*	−.050
	(.129)	(.142)	(.142)	(.330)	(.230)
LC today		.132	.138	−.920	.068
		(.286)	(.287)	(.820)	(.539)
Male	.106	.110	.101	.098	.278*
	(.068)	(.068)	(.069)	(.071)	(.143)
Age	.209	.136	.183	−.435	.390
	(.316)	(.334)	(.344)	(.540)	(.688)
Experience	.073	.073	.068	.087*	.002
	(.047)	(.047)	(.048)	(.051)	(.089)
Parent entrepreneur	−.088	−.086	−.086	−.116	.121
	(.069)	(.070)	(.071)	(.075)	(.160)
Married	.117	.118	.113	.087	.228
	(.090)	(.090)	(.091)	(.099)	(.183)
Family size	−.006	−.006	−.002	−.018	−.013
	(.033)	(.033)	(.033)	(.035)	(.057)
GDP growth at learning age	−.002	−.138	.152	1.292	10.853
	(4.471)	(4.490)	(4.562)	(5.603)	(8.378)
ED birth			5.478		
			(4.380)		
λ_{mover}					.432
					(.421)
Observations	2,627	2,627	2,588	2,627	667
Area dummies:					
Macro area of residence	X	X	X		X
Macro area of learning	X	X	X		X
Province of residence				X	
Province of learning				X	

NOTE.—The dependent variable is log of sales per worker. All regressions are based on data for the years 2005–7 and include year, education, and sector dummies. Column 5 uses only the sample of movers, correcting for selection with a Heckman model in which the excluded variable is the average mobility rate out of the municipality where the individual was living at learning age in the 10 years after the learning age. Standard errors clustered at the local labor system and year of learning level are in parenthesis. GDP = gross domestic product.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Table A4
Entrepreneurial Traits and Total Factor Productivity

	(1)	(2)	(3)	(4)	(5)
Risk aversion	−.035				
	(.042)				
Ambiguity aversion		.013			
		(.020)			
Confidence			.096		
			(.074)		
Optimism				.015	
				(.016)	
Satisfaction					.020
					(.018)
ED learn	1.395	1.061	1.534	1.449	1.820
	(1.945)	(1.961)	(1.979)	(1.973)	(1.991)
ED today	−2.010	−1.768	−2.157	−2.364	−2.488
	(2.654)	(2.668)	(2.694)	(2.672)	(2.706)
Male	−.069	−.078	−.087	−.078	−.072
	(.063)	(.063)	(.064)	(.064)	(.064)
Experience	.035	.027	.040	.040	.044
	(.037)	(.037)	(.038)	(.038)	(.038)
Parent entrepreneur	−.095	−.092	−.101	−.105*	−.110*
	(.061)	(.061)	(.062)	(.061)	(.062)
Junior/high school	.183	.183	.169	.185	.187
	(.112)	(.113)	(.113)	(.113)	(.115)
College/postgraduate	−.004	−.001	−.025	−.007	.017
	(.069)	(.069)	(.070)	(.069)	(.070)
Observations	967	955	944	946	937

NOTE.—All regressions include year, education, macro area of learning and of current location, and sector dummies.
* Significant at 10%.

Table A5
Managerial Practices, Total Factor Productivity (TFP), and Firm Size

	TFP				Employment (Log)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Management	.152**				.607***			
	(.067)				(.073)			
Monitoring		.100*				.482***		
		(.052)				(.057)		
Targets			.074				.346***	
			(.049)				(.056)	
People				.191***				.552***
				(.068)				(.076)
ED learn	−2.126	−2.012	−1.754	−2.277	.160	.095	1.394	.567
	(2.974)	(2.980)	(2.978)	(2.961)	(3.255)	(3.245)	(3.367)	(3.313)
ED today	.119	.225	.275	−.250	−3.469	−2.913	−2.705	−4.608
	(3.924)	(3.932)	(3.942)	(3.910)	(4.295)	(4.282)	(4.456)	(4.374)
Male	−.202**	−.192*	−.181*	−.204**	.056	.072	.126	.097
	(.098)	(.098)	(.097)	(.097)	(.107)	(.106)	(.110)	(.109)
Experience	.083	.080	.079	.073	.092	.094	.088	.046
	(.070)	(.070)	(.070)	(.069)	(.076)	(.076)	(.079)	(.077)
Parent entrepreneur	−.174**	−.183**	−.182**	−.163*	.116	.089	.093	.131
	(.088)	(.088)	(.088)	(.088)	(.096)	(.095)	(.099)	(.098)
Junior/high school	.166	.147	.159	.166	−.009	−.075	−.022	−.044
	(.187)	(.187)	(.188)	(.186)	(.205)	(.204)	(.213)	(.208)
College/postgraduate	.039	.053	.046	.026	.071	.114	.086	.066
	(.098)	(.098)	(.098)	(.098)	(.107)	(.106)	(.111)	(.109)
Observations	386	386	386	386	386	386	386	386

Note.—All regressions include year, education, macro area of learning and of current location, and sector dummies.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

References

Åstebro, Thomas. 2003. The return to independent invention: Evidence of unrealistic optimism, risk seeking or skewness loving? *Economic Journal* 113, no. 484:226–39.

Banerjee, Abhijit V., and Andrew Newman. 1994. Poverty, incentives and development. *American Economic Review Papers and Proceedings* 84:211–15.

Barlow, Richard E., and Frank Proschan. 1975. *Statistical theory of reliability and life testing*. New York: Holt, Rinehart & Winston.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. 2018. Who becomes an inventor in America? The importance of exposure to innovation. *Quarterly Journal of Economics* 134, no. 2:647–713.

Bisin, Alberto, and Thierry Verdier. 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory* 97:298–319.

Bloom, Nicholas, Christos Genakos, Raffaella Sadun, and John Van Reenen. 2012a. Management practices across firms and countries. *Academy of Management Perspectives* 26:12–33.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2012b. Americans do I.T. better: US multinationals and the productivity miracle. *American Economic Review* 102:167–201.

Bloom, Nicholas, and John Van Reenen. 2007. Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics* 122, no. 4:1351–408.

———. 2010a. New approaches to surveying organizations. *American Economic Review* 100, no. 2:105–9.

———. 2010b. Why do management practices differ across firms and countries? *Journal of Economic Perspectives* 24, no. 1:203–24.

Camerer, Colin, and Dan Lovallo. 1999. Overconfidence and excess entry: An experimental approach. *American Economic Review* 89, no. 1:306–18.

Cesarini, David, Christopher T. Dawes, Magnus Johannesson, Paul Lichtenstein, and Björn Wallace. 2009. Genetic variation in preferences for giving and risk taking. *Quarterly Journal of Economics* 124:809–42.

Chinitz, Benjamin. 1961. Contrasts in agglomeration: New York and Pittsburgh. *American Economic Review* 51:279–89.

Cingano, Federico, and Fabiano Schivardi. 2004. Identifying the sources of local productivity growth. *Journal of the European Economic Association* 2:720–42.

De Figueiredo, Rui J. P., Philipp Meyer-Doyle, and Evan Rawley. 2013. Inherited agglomeration effects in hedge fund spawns. *Strategic Management Journal* 34, no. 7:843–62.

De La Roca, Jorge, and Diego Puga. 2017. Learning by working in big cities. *Review of Economic Studies* 84:106–42.
Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2012. The intergenerational transmission of risk and trust attitudes. *Review of Economic Studies* 79:645–77.
Duranton, Gilles, and Diego Puga. 2004. Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics*, vol. 4, ed. J. Vernon Henderson and Jacques-François Thisse, 2063–117. Amsterdam: Elsevier.
Erikson, Erik H. 1968. *Identity, youth and crisis*. New York: Norton.
Evans, David S., and Boyan Jovanovic. 1989. An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy* 97:808–27.
Gendron-Carrier, Nicolas. 2018. Understanding the careers of young entrepreneurs. Unpublished manuscript.
Glaeser, Edward L., Hedi D. Kallal, José A. Scheinkman, and Andrei Shleifer. 1992. Growth in cities. *Journal of Political Economy* 100:1125–52.
Glaeser, Edward L., William R. Kerr, and Giacomo A. M. Ponzetto. 2010. Clusters of entrepreneurship. *Journal of Urban Economics* 67:150–68.
Glaeser, Edward L., and David C. Maré. 2001. Cities and skills. *Journal of Labor Economics* 19:316–42.
Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2004. Does local financial development matter? *Quarterly Journal of Economics* 119:929–69.
Guiso, Luigi, and Fabiano Schivardi. 2011. What determines entrepreneurial clusters? *Journal of the European Economic Association* 9, no. 1:61–86.
Harris, Judith Rich. 2011. *The nurture assumption: Why children turn out the way they do*. New York: Free Press.
Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review* 103:2052–86.
Hurrelmann, Klaus. 1988. *Social structure and personality development: The individual as a productive processor of reality*. Cambridge: Cambridge University Press.
Jacobs, Jane. 1969. *The economy of cities*. New York: Vintage.
Kihlstrom, Richard, and Jean-Jacques Laffont. 1979. A general equilibrium theory of firm formation based on risk aversion. *Journal of Political Economy* 87:719–48.
Knight, Frank H. 1921. Risk, uncertainty and profit. New York: Houghton Mifflin.
Lazear, Edward P. 2005. Entrepreneurship. *Journal of Labor Economics* 23:649–80.

Lindquist, Matthew J., Joeri Sol, and Mirjam Van Praag. 2015. Why do entrepreneurial parents have entrepreneurial children? *Journal of Labor Economics* 33:269–296.
Lucas, Robert E., Jr. 1978. On the size distribution of business firms. *Bell Journal of Economics* 2:508–23.
Manski, Charles. 1994. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies* 60:531–42.
Marshall, Alfred. 1890. *Principles of economics*. London: Macmillan.
Matano, Alessia, and Paolo Naticchioni. 2016. What drives the urban wage premium? Evidence along the wage distribution. *Journal of Regional Science* 56:191–209.
McFadden, Daniel L. 1974. Conditional logit analysis of qualitative choice behavior. In *Frontiers in econometrics*, ed. P. Zarembaka, 105–42. New York: Academic Press.
Michelacci, Claudio, and Olmo Silva. 2007. Why so many local entrepreneurs? *Review of Economics and Statistics* 89, no. 4:615–33.
Moretti, Enrico. 2011. Local labor markets. In *Handbook of labor economics*, vol. 4, ed. D. Card and O. Ashenfelter, 1237–313. Amsterdam: Elsevier.
Nicolaou, Nicos, and Scott Shane. 2010. Entrepreneurship and occupational choice: Genetic and environmental influences. *Journal of Economic Behavior and Organization* 76:3–14.
Nicolaou, Nicos, Scott Shane, Lynn Cherkas, Janice Hunkin, and Tim D. Spector. 2008. Is the tendency to engage in entrepreneurship genetic? *Management Science* 54:167–79.
Rosen, Sherwin. 1982. Authority, control, and the distribution of earnings. *Bell Journal of Economics* 13:311–23.
Rosenthal, Stuart S., and William C. Strange. 2004. Evidence on the nature and sources of agglomeration economies. In *Handbook of regional and urban economics*, vol. 4, ed. J. Vernon Henderson and Jacques-François Thisse, 2119–2171. Amsterdam: Elsevier.
Scheier, Michael F., Charles S. Carver, and Michael W. Bridges. 1994. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the life orientation test. *Journal of Personality and Social Psychology* 67, no. 6:1063–78.
Schumpeter, Joseph. 1911. *The theory of economic development*. Cambridge, MA: Harvard University Press.
Williamson, Oliver E. 2000. The new institutional economics: Taking stock, looking ahead. *Journal of Economic Literature* 38:595–613.
Zhang, Zhen, Michael J. Zyphur, Jayanth Narayanan, Richard D. Arvey, Sankalp Chaturvedi, Bruce J. Avolio, Paul Lichtenstein, and Gerry Larsson. 2009. The genetic basis of entrepreneurship: Effects of gender and personality. *Organizational Behavior and Human Decision Processes* 110, no. 2:93–107.