SMART AND ILLICIT: WHO BECOMES AN ENTREPRENEUR AND DO THEY EARN MORE?*

Ross Levine and Yona Rubinstein

We disaggregate the self-employed into incorporated and unincorporated to distinguish between "entrepreneurs" and other business owners. We show that the incorporated self-employed and their businesses engage in activities that demand comparatively strong nonroutine cognitive abilities, while the unincorporated and their firms perform tasks demanding relatively strong manual skills. People who become incorporated business owners tend to be more educated and—as teenagers—score higher on learning aptitude tests, exhibit greater self-esteem, and engage in more illicit activities than others. The combination of "smart" and "illicit" tendencies as youths accounts for both entry into entrepreneurship and the comparative earnings of entrepreneurs. Individuals tend to experience a material increase in earnings when becoming entrepreneurs, and this increase occurs at each decile of the distribution. *JEL Codes*: L26, J24, J3, G32.

I. Introduction

Economists since Adam Smith (1776) have emphasized that entrepreneurs spur improvements in living standards. For example, Schumpeter (1911) argues that entrepreneurs drive economic growth by undertaking risky ventures that create and introduce new goods, services, and production processes that displace old businesses. Lucas (1978), Baumol (1990), Murphy, Shleifer, and Vishny (1991), and Gennaioli et al. (2013) stress that the human capital of entrepreneurs plays a unique role in shaping the productivity of firms and the growth rate of entire economies.

Yet a substantial body of research—using data on the self-employed to draw inferences about entrepreneurship—concludes that the median entrepreneur earns less than a salaried counterpart (e.g., Borjas and Bronars 1989; Evans and Leighton 1989; Hamilton 2000; Moskowitz and Vissing-Jørgensen 2002). Furthermore, as we document below, the self-employed and salaried workers have similar education, learning aptitude scores, and family backgrounds. If the self-employed are a good proxy for risk-taking, growth-creating entrepreneurs, it is puzzling that their human capital traits are similar to those of salaried workers and that they earn less.

Perhaps, self-employment is not a good proxy for entrepreneurship. Glaeser (2007) argues that self-employment aggregates together different types of activities and individuals, making "little distinction between Michael Bloomberg and a hot dog vendor." In a cross-section of developing economies, La Porta and Shleifer (2008, 2014) disaggregate the self-employed into those running formal or informal firms and find that informal firms tend to be low-productivity businesses run by poorly educated owners, while formal firms tend to be higher-productivity enterprises with more educated owners. For the United States, Evans and Leighton (1989) find that although some of the self-employed are productivity-enhancing entrepreneurs, most are one-person retail business owners who did not succeed as salaried workers, and Hurst and Pugsley (2011) show that only a few of the self-employed seek to grow. Thus, studying the self-employed in general might yield misguided inferences about entrepreneurs in particular.

In this article we offer a different empirical proxy for entrepreneurship and use it to assess who becomes an entrepreneur and whether they earn more. In parallel with La Porta and Shleifer's (2008, 2014) differentiation between formal and informal firms in developing economies, we disaggregate the self-employed into the incorporated and unincorporated to distinguish between "entrepreneurs" and other business owners in the United States. This is a natural disaggregation given the costs and benefits associated with incorporation. Although incorporated business owners face additional fees and regulations, they benefit from the corporation's key legal features—limited liability and a separate legal identity—that limit the financial and legal risk of the owners. These legal features are especially valuable to business owners seeking to undertake large, risky investments. In

contrast, people will tend to choose the unincorporated legal form when these features are less important.

Consistent with using the incorporated as a better proxy for entrepreneurship than the aggregate group of self-employed, we discover that the incorporated tend to engage in activities and open businesses that are more closely aligned with core conceptions of entrepreneurship than the unincorporated. Specifically, using the U.S. Department of Labor's Dictionary of Occupational Titles, we show that the incorporated—and their businesses— perform activities demanding comparatively strong nonroutine cognitive skills, such as (i) creativity, analytical flexibility, and generalized problem solving and (ii) complex interpersonal communications associated with persuading and managing. We view these activities as closely aligned with productivity-enhancing entrepreneurship. In contrast, the unincorporated tend to engage in activities that demand relatively low levels of these cognitive skills and high levels of eye, hand, and foot coordination, for example, landscaping, truck driving, and carpentry. Furthermore, we find that unincorporated businesses rarely incorporate and incorporated ones rarely become unincorporated sole proprietorships or partnerships. These results suggest that the choice of the business's legal form largely reflects the ex ante nature of the business, not its ex post performance.

We next turn to the question of who becomes an entrepreneur. Using data from the Current Population Survey (CPS) and the National Longitudinal Survey of Youth (NLSY), we show that those who become incorporated business owners have distinct cognitive and noncognitive traits. The incorporated tend to be more educated and—as teenagers—score higher on learning aptitude tests, exhibit greater self-esteem, reveal stronger sentiments of controlling their futures, and engage in more illicit activities than others. Moreover, it is a particular mixture of early determined traits that is most powerfully associated with incorporated self-employment. People who both participated in illicit activities as teenagers and scored highly on learning aptitude tests as youths— the "smart and illicit"—have a much greater tendency to become incorporated business owners than others.

With respect to earnings, we discover the following. First, on average and at the median, the incorporated self-employed earn more than comparable salaried workers, while the unincorporated self-employed earn much less. Second, much of this earnings gap reflects person-specific influences: incorporated

business owners were typically highly successful salaried workers even after accounting for Mincerian skills, while the unincorporated were comparatively unsuccessful employees. Third, despite the person-specific effects, people who self-select into incorporated self-employment tend to experience a sizeable increase in earnings beyond their high incomes as salaried workers, while an individual's earnings tend to fall when he switches from a salaried job to unincorporated self-employment. Fourth, the increase in earnings associated with becoming an incorporated business owner does not reflect selection into incorporated self-employment on ex post success as an unincorporated business owner, that is, few people start as unincorporated business owners and then incorporate; and, accounting for these few switchers does not materially alter our findings. Fifth, the finding that earnings tend to rise when individuals become incorporated self-employed holds at each decile of the distribution when accounting for person-specific effects. Sixth, the incorporated self-employed have a much wider dispersion of earnings than those of salaried workers, but this dispersion becomes much tighter when accounting for person-specific effects.

We also show that many of the same traits that explain selection into entrepreneurship also account for success as an entrepreneur. People with both high learning aptitude and high illicit activity scores as youths tend to experience much larger increases in earnings when they become incorporated self-employed business owners than people without that combination of traits. Yet this combination of "smart and illicit" traits is associated with smaller earnings for unincorporated business owners. While past research shows the importance of noncognitive traits for labor market outcomes (e.g., Heckman 2000; Bowles, Gintis, and Osborne 2001; Heckman and Rubinstein 2001; Heckman, Stixrud, and Urzua 2006), we document that some mixtures of traits receive positive or negative remuneration depending on the activity.

Our findings stress that better educated people with "smart and illicit" traits as youths are more likely than others to establish incorporated businesses and experience a material increase in earnings, while people without this constellation of human capital traits who become self-employed tend to run relatively unsuccessful, unincorporated businesses and experience a drop in earnings. Since neither selection into self-employment nor the subsequent success of the business is exogenous to person traits, our findings

do not reflect the causal impact of randomly making somebody incorporated or unincorporated on earnings. Rather, the findings reflect the change in earnings associated with a person choosing to run an incorporated or unincorporated business.

Our work relates to research seeking to explain why the median self-employed business owner earns less per hour than a comparable salaried employee. To explain this finding, Hurst and Pugsley (2011) show that the nonpecuniary benefits of self-employment, such as being "one's own boss," help attract people into self-employment even when earnings are lower. De Meza and Southey (1996), Bernardo and Welch (2001), and Dawson et al. (2014) stress that the "overconfidence" of the self-employed could explain both entry into self-employment and their comparatively low earnings. We offer a different, though not mutually exclusive, explanation. We emphasize that self-employment is a poor proxy for entrepreneurship and show that incorporated self-employment is better. The incorporated tend to be "smart and illicit," have strong nonroutine cognitive skills, open businesses demanding these same cognitive abilities from their workers, and enjoy a sharp increase in earnings when becoming business owners. In turn, the unincorporated tend to have average learning aptitude scores, weak nonroutine cognitive skills, strong manual skills, open one-person businesses, and experience a drop in hourly earnings when becoming self-employed. Since there are more unincorporated than incorporated, this accounts for earlier findings on the negative pecuniary returns to self-employment.

The article is organized as following. Section II presents the data and summary statistics. Section III examines the different job task requirements of incorporated and unincorporated business owners and their employees. We study who becomes an entrepreneur and whether they earn more in Sections IV and V, respectively. Section VI concludes.

II. The Data and Summary Statistics

II.A. CPS

We use the March Annual Demographic Survey files of the CPS for the work years 1995 through 2012.[1] We start in 1995 because (i) the measure of incorporation changed following the

---

1. The Online Data Appendix provides details on the construction of all data used in this paper, including our use of data from the CPS, NLSY79, and Dictionary

redesign of the CPS in 1994 (Hipple 2010), (ii) the CPS improved its top-coding in work year 1995 by allowing for differences across classes of workers and demographics, and (iii) the post-1995 period corresponds closely to the relevant years from the NLSY79. In "our" main CPS sample, we include prime age workers (25–55 years old) and exclude (i) people with missing data on age, race, gender, schooling, industry codes, or occupation codes, and (ii) those living in group quarters or working in agriculture or the military.

The CPS classifies all workers in each year as either salaried or self-employed and, among the self-employed, indicates whether individuals are incorporated or unincorporated. Specifically, individuals are asked about their employment class for their main job: "Were you employed by a government, by a private company, a nonprofit organization, or were you self-employed (or working in a family business)?" Those responding that they are self-employed are further asked, "Is this business incorporated?" While incorporation offers the benefits of limited liability and a separate legal identity, there are direct costs of incorporation, such as annual fees and the preparation of more elaborate financial statements, and indirect costs associated with the separation of ownership and control. In terms of occupation, about half of the incorporated self-employed are managers and no other three-digit occupation accounts for more than 3.5% of the incorporated self-employed. Physicians and surgeons (3.3%), lawyers (3.3%), and accountants (1.3%) combine to account for less than 8% of incorporated self-employment.[2] With respect to the unincorporated, about 25% are managers. Carpenters (9.2%), truck drivers (4.6%), and automobile mechanics (3.5%) combine to account for about 17% of unincorporated self-employment.

We also construct a two-year matched panel. The CPS interviews a household for four consecutive months. The next year, the CPS returns to the same location. In most cases, the second interview involves the same household as the first interview. We follow the guidelines in Madrian and Lefren (2000) for matching

---

of Occupational Titles. The Online Appendix Tables I, II, IV, VII.A, VII.B, VIII, and XI show that the results are robust to small perturbations in the construction of the main variables.

2. The results are robust both to excluding these occupations or including separate dummy variables for physicians and surgeons, lawyers, and accountants as shown in the Online Appendix Table IX.C.

CPS households across time. This involves checking the age, race, gender, education, and so on of those interviewed.[3]

Table I provides summary statistics from the CPS on the age, race, gender, education, and labor market outcomes of individuals reported as working while distinguishing among salaried workers, all self-employed workers, the unincorporated self-employed, and the incorporated self-employed. Hourly earnings are defined as real annual earnings divided by the product of weekly working hours and annual working weeks, where the Consumer Price Index is used to deflate earnings to 2010 dollars. All CPS calculations are weighted using the March supplement weights.

Compared to the median self-employed individual, the median salaried worker earns more per hour and has similar educational attainment. For example, salaried workers have on average 13.7 years of education, while the self-employed have 13.9. These summary statistics confirm the puzzle emerging from the extant literature: if entrepreneurship drives technological innovation and growth, it is odd that the self-employed, which are often used to draw inferences about entrepreneurship, earn less, and have similar levels of education as salaried workers.

The demarcation between the incorporated and unincorporated highlights two differences. First, the median incorporated individual earns much more per hour—and works many more hours—than the median salaried and unincorporated individual. Indeed, median hourly earnings of the incorporated are about 80% greater than that of the unincorporated and 35% more than the median salaried employee. Second, the incorporated self-employed have distinct demographic and educational traits. The incorporated tend to be disproportionately white, male, and highly educated. For example, women account for 48% of the sample of workers but only 28% of the incorporated. As another example, while 33% of salaried workers graduate from college, 46% of the incorporated have a college degree.[4] Simply comparing salaried

3. We find no evidence for differential selection into the matched-CPS sample once we condition on demographics, as discussed in the Online Data Appendix and shown in Online Appendix Table IX.B.

4. Research by Bertrand and Schoar (2003), Bloom and Van Reenen (2007), Malmendier and Tate (2009), and Queiro (2015) shows that the human capital of senior managers influences firm-level productivity. We focus on business owners, not managers, and examine how education and distinct cognitive and noncognitive traits influence entry into entrepreneurship and success as a business owner.

TABLE I

DEMOGRAPHICS AND LABOR MARKET OUTCOMES BY EMPLOYMENT TYPE

			Self-employed		
	All	Salaried	All	Unincorporated	Incorporated
Panel A: CPS 1996–2012					
Observations	1,225,886	1,108,591	117,295	75,476	41,819
	100.0%	90.4%	9.6%	6.2%	3.4%
A. Labor market outcomes					
Mean earnings	$ 47,515	$ 46,421	$ 58,174	$ 40,820	$ 89,169
Median earnings	$ 36,090	$ 36,363	$ 34,190	$ 24,625	$ 55,591
Median hourly earnings	$ 18.0	$ 18.0	$ 17.4	$ 13.8	$ 24.6
Annual hours worked	1,985	1,976	2,078	1,936	2,331
Full-time, full-year	0.69	0.70	0.64	0.57	0.78
B. Demographics					
Age	40.2	40.0	42.9	42.4	43.6
White	0.70	0.69	0.79	0.76	0.83
Female	0.48	0.49	0.36	0.40	0.28
Years of schooling	13.7	13.7	13.9	13.6	14.5
College graduate (or more)	0.33	0.33	0.36	0.31	0.46
Panel B: NLSY79 1982–2012					
Observations	132,681	121,782	10,899	8,963	1,936
	100.0%	91.8%	8.2%	6.8%	1.5%
A. Labor market outcomes					
Mean earnings	$ 44,725	$ 43,605	$ 55,785	$ 45,713	$ 93,411
Median earnings	$ 35,170	$ 35,222	$ 33,965	$ 28,672	$ 61,424
Median hourly earnings	$ 17.2	$ 17.2	$ 16.8	$ 14.7	$ 26.2
Annual hours worked	1,966	1,953	2,088	1,991	2,461
Full-time, full-year	0.59	0.59	0.53	0.48	0.72
B. Demographics					
Age	36.2	36.0	38.1	37.5	40.1
White	0.81	0.80	0.87	0.86	0.90
Female	0.47	0.48	0.38	0.41	0.28
Years of schooling	13.8	13.8	13.6	13.4	14.2
College graduate (or more)	0.30	0.30	0.26	0.23	0.36
C. Firm size: number of employees					
Median			0.0	0.0	2.0
Mean			8.6	2.1	23.0

Notes. The table presents summary statistics from the March Annual Demographic Survey files of the Census Bureau's CPS for the work years 1995 through 2012, for prime age workers (25 through 55 years old), and from the Bureau of Labor Statistics' National Longitudinal Survey of Youth 1979 (NLSY79) for workers who are least 25 years old between 1982 and 2012. The CPS and the NLSY79 classify workers in each year as either salaried or self-employed, and among the self-employed, they indicate whether the person is incorporated or unincorporated self-employed. The number of employees includes all paid employees in the year that the person becomes full-time self-employed and excludes the self-employed business owner, which is available from 2002 onward in the NLSY79. When using the CPS, we further exclude observations with missing data on age, race, gender, schooling, industry codes, or occupation codes, and those living in group quarters or working in agriculture or the military. When using the NLSY79, we further exclude observations with missing values on age, race, or cognitive and noncognitive traits (AFQT, Rosenberg Self-Esteem and Rotter Locus of Control). The Online Data Appendix provides further details on the sample and variables.

and self-employed workers conceals huge differences across employment types.[5]

II.B. NLSY79: Basics

The NLSY79 is a representative survey of 12,686 individuals who were 15–22 years old when they were first surveyed in 1979.[6] Individuals were surveyed annually through 1994 and have since been surveyed biennially. To examine individuals who are 25 years of age or older, we use survey years starting in 1982. Since nobody in our sample is above the age of 55 in the last survey year of our sample (2012), this NLSY79 subsample corresponds to that of the CPS analyses. From this subsample, our "main" NLSY79 sample includes salaried and self-employed individuals and excludes individuals with missing data on age, gender, race, or cognitive and noncognitive traits reported by the NLSY79 (AFQT, Rosenberg Self-Esteem and Rotter Locus of Control), which we define below. Since the NLSY79 survey is conducted every other year after 1994, we use year $t - 2$ when examining lagged values of respondent characteristics for all years in the NLSY79 analyses.

Although the NLSY79 surveys a smaller cross section of people than the CPS, it has two advantages. First, the NLSY79 is an extensive panel that traces individuals from when they were 15–22 years old through the age of 48–55. Thus, we follow virtually the entire career path of individuals. Second, the NLSY79 provides detailed information about the cognitive and noncognitive traits of individuals before they become prime age workers. Thus, we can examine how the traits of individuals when

5. As noted in the Introduction, we are not the first to recognize that self-employment is a weak proxy for entrepreneurship. For example, La Porta and Shleifer (2008, 2014) disaggregate the self-employed into those running formal and informal firms to differentiate between different types of self-employment. Others also stress that self-employment aggregates together low- and high-productivity businesses, notably Evans and Leighton (1989), Carr (1996), Budig (2006), Mohapatra, Rozelle, and Goodhue (2007), Hurst and Pugsley (2011), Özcan (2011), and Bengtsson, Sanandaji, and Johannesson (2012), and offer strategies for differentiating between such enterprises.

6. We use the cross-sectional (6,111 individuals), the supplemental (5,295 individuals), and military (1,280 individuals) samples. Also, since the NLSY draws on a slightly younger sample of individuals and the incorporated self-employed are older than other employment types, a smaller percentage of the NLSY sample is incorporated than in the CPS. For comparisons of the CPS and NLSY, see Fairlie (2005) and Fairlie and Meyer (1996).

they were teenagers account for career choices later in life. We describe these unique traits in the next subsection.

The Table I summary statistics from the NLSY79 and CPS provide similar messages about labor market outcomes and demographics.[7] First, the median earnings of salaried workers are greater than those of the self-employed. Second, this conceals enormous differences between the incorporated and unincorporated self-employed. The median incorporated individual earns about 50% more per hour and works about 25% more hours than the median salaried worker. In contrast, the median unincorporated business owner earns about 15% less per hour than the median salaried worker. Third, the incorporated are disproportionately white, male, and highly educated, while the unincorporated tend to be less educated than salaried workers. The incorporated are notably different from the unincorporated. Hurst, Li, and Pugsley (2014) show that the self-employed underreport their incomes, which might account for some of the lower median earnings reported by the unincorporated. Finally, incorporated firms have, on average and at the median, more employees than unincorporated ones. We examine the number of employees in the year that a person becomes a full-time business owner. As shown, the incorporated average 23 employees, while the unincorporated average 2. At the median, the incorporated business has two employees, while the unincorporated has no paid employees and is a one-person business.[8]

7. With respect to the NLSY79, we measure educational attainment at the end of the respondent's educational experience, so it does not vary over time for a respondent. Furthermore, since the unit of analysis is an individual-year and some people work in different employment types during their careers, we weight by the number of years the person worked in each type when providing summary statistics in Table I about fixed characteristics by employment type.

8. We therefore use firm size differences in our analyses below. In these summary statistics we focus on the NLSY79 because the CPS only provides statistics on employment "bins," where the smallest bin size includes firms with fewer than 10 workers and where 84% of all firms fall into this bin size. Of the unincorporated self-employed, 92% have fewer than 10 workers, while the comparable figure for the incorporated is 72%.

II.C. NLSY79: Cognitive and Noncognitive Traits

The NLSY79 also provides the following information on individual and family traits.

AFQT score (Armed Forces Qualifications Test score) measures the aptitude and trainability of each individual. Collected during the 1980 NLSY79 survey, the AFQT score is based on arithmetic reasoning, world knowledge, paragraph comprehension, and numerical operations. It is frequently employed as a general indicator of cognitive skills. This AFQT score is measured as a percentile of the NLSY79 survey, with a median value of 50.

Rosenberg Self-Esteem score, which is based on a ten-part questionnaire given to all NLSY79 participants in 1980, measures the degree of approval or disapproval of one's self and has been widely used in psychology and economics (e.g., Bowles, Gintis, and Osborne 2001; Heckman, Stixrud, and Urzua 2006). The values range from 6 to 30, where higher values signify greater self-approval.

Rotter Locus of Control measures the degree to which individuals believe they have internal control of their lives through self-determination relative to the degree that external factors, such as chance, fate, and luck, shape their lives. It was collected as part of a psychometric test in the 1979 NLSY79 survey. The Rotter Locus of Control ranges from 4 to 16, where higher values signify less internal control and more external control.

Illicit Activity Index measures the aggressive, risk taking, disruptive, "break-the-rules," behavior of individuals based on the 1980 survey. We construct this index from 20 questions, where 17 are questions about delinquency and three are about interactions with the police. The delinquency questions cover issues associated with damaging property, fighting at school, shoplifting, robbery, using force to obtain things, assault, threatening to assault somebody, drug use, dealing drugs, gambling, and so forth. The police questions involve being stopped by the police, charged with an illegal activity, or convicted, all for activities other than minor traffic offenses. For each question, we assign the value 1 if the person responds in 1980 the he or she engaged in that activity and 0 otherwise. To obtain the index, we simply add these values and divide by 20. To construct the standardized version we then subtract the sample mean and divide by the standard deviation,

so that the Illicit Activity Index (standardized) has a mean of 0 and a standard deviation of 1. Higher values signify more illicit behaviors.[9] We also report results using the answers to some of the individual questions, such as whether the person used force to obtain things (Force), stole something of $50 or less (Steal 50 or less), and whether the person was Stopped by the Police.[10]

While some might view the Illicit Activity Index as only proxying (inversely) for risk aversion, our analyses caution against this presumption and hence highlight the degree to which the Illicit Activity Index measures the aggressive, disruptive activities of individuals as youths. After controlling for other traits, there is not a strong association between the Illicit Activity Index (measured in 1980) and the NLSY79's risk aversion indicator that assesses for how much a person would sell an item that has an expected, though risky, future value of $5,000 (measured in 2006).

We use additional information on each individual's pre-labor market family traits. Family Income in 1979 equals the respondent's family income in 1979 in 2010-year prices. In those cases where 1979 is missing, we use the earliest year between 1980 and 1981 with a nonmissing value. Father's Education and Mother's Education equal the years of schooling of the respondent's father and mother respectively. Two Parent Family (14) equals 1 if the respondent lived in a two-parent family at the age of 14 and 0 otherwise. For additional details on these pre–labor market family traits, and the other variables used in the paper, see the Online Data Appendix.

The NLSY79 also posed new questions in 2010 that provide helpful information in assessing the validity of using the unincorporated and incorporated self-employed as indicators of the ex ante nature of the business venture. To measure the degree to

9. As described in the Online Data Appendix and shown in Online Appendix Tables VII.A, VIII, and XI, the results are robust to computing the Illicit Activity Index in different ways. In particular, to some of the questions composing the index, the survey offers yes or no answers and to other questions it offers answers that yield information on the frequency with which the person engages in a "delinquent" activity. In the article, we code the responses to all questions as 1 or 0 based on whether the respondent did or did not engage in any of the activity. However, we show that all of the results hold when using information on the frequency with which individuals engage in delinquent activities.

10. As a caveat, note that the NLSY79 is a survey, so the Illicit Activity Index is based on each person's willingness to report their engagement in illicit activities. For details on this index, see the Online Data Appendix.

which an individual considers himself to be an entrepreneur, we use Entrepreneur, which equals 1 if the respondent in 2010 answers "yes" to the question, "Do you consider yourself to be an entrepreneur?" In posing the question, the NLSY79 defines an entrepreneur as "someone who launches a business enterprise, usually with considerable initiative and risk." To provide some information on the degree to which the individual is engaged in an innovative activity, we use Applied for Patent, which equals 1 if the respondent in 2010 answered, "yes" to the question, "Has anyone, including yourself, ever applied for a patent for work that you significantly contributed to?"

Individuals who become incorporated self-employed have distinct family backgrounds, as shown in Panel A of Table II, which uses the main NLSY79 sample from Table I. Compared to others, the incorporated self-employed come from (i) high-income families as measured by Family Income in 1979, (ii) well-educated families as measured by the education of the respondent's parents, and (iii) two-parent families as measured by Two Parent Family (14).

Moreover, individuals who become incorporated self-employed display striking cognitive and noncognitive characteristics before they enter the labor market (Table II, Panel B).[11] First, people who become incorporated self-employed had (i) higher "ability" as measured by AFQT values, (ii) stronger self-esteem as measured by Rosenberg scores, and (iii) stronger senses of controlling their futures, rather than having their futures determined by fate or luck, as measured by low Rotter Locus of Control scores. Second, people who spend more of their prime age working years as incorporated engaged in more illicit activities as youths. For example, the incorporated are twice as likely as salaried workers to report having taken something by force as youths; they are almost 40% more likely to have been stopped by the police; and the incorporated self-employed have an overall illicit activity index (measured when they were between the ages of 15 and 22 and standardized for the full sample) that is 0.2 standard deviations greater than that of salaried workers. Furthermore, while the unincorporated also tend to engage in more illicit activities as

11. We report standardized values, so Rotter Locus of Control (standardized), Rosenberg Self-Esteem (standardized), and Illicit Activity Index (standardized) each has a mean of 0 and a standard deviation within the full NLSY79 sample.

TABLE II

HOME ENVIRONMENT, EARLY PERSONAL TRAITS, AND OTHER CHARACTERISTICS

			Self-employed		
	All	Salaried	All	Unincorporated	Incorporated
Panel A: Family background					
Mother's education	11.7	11.7	12.0	11.8	12.6
Father's education	11.9	11.9	12.2	12.1	12.7
Two-parent family (14)	0.76	0.76	0.77	0.76	0.83
Family income in 1979	$58,185	$57,894	$60,940	$58,246	$71,384
Panel B: Cognitive and noncognitive traits					
AFQT	50.1	50.0	51.4	50.4	55.2
Rotter locus of control (stand.)	-0.10	-0.09	-0.18	-0.16	-0.28
Rosenberg self-esteem (stand.)	0.08	0.07	0.10	0.06	0.27
Illicit activity index (stand.)	0.01	0.00	0.12	0.10	0.20
Force (raw)	0.04	0.04	0.06	0.06	0.08
Steal 50 or less (raw)	0.21	0.21	0.24	0.23	0.26
Stopped by police (raw)	0.19	0.18	0.22	0.21	0.26
Panel C: Self-designation and invention					
Entrepreneur (residual stand.)	0.00	−0.08	0.80	0.69	1.20
Applied for patent (residual stand.)	0.00	−0.01	0.08	0.03	0.28

Notes. This table provides summary statistics from the NLSY79 on people who are at least 25 years old and in the work force, for sample years 1982 through 2012, as in Table I. Family background and data on cognitive and noncognitive traits are measured in 1979 and 1980, which is before anyone in the sample enters prime age. Mother's Education and Father's Education are the number of years of education of the person's mother and father. Two-Parent Family (14) equals 1 if the person at the age of 14 had two parents living at home and 0 otherwise. Family Income in 1979 is the income of the person's family in 1979, measured in 2010-year prices. When Family Income is missing in 1979, we use the earliest year between 1980 and 1981 with a nonmissing value. AFQT is a measure of the aptitude and trainability of the respondent; Rotter Locus of Control measures the degree to which respondents believe they have internal control of their lives through self-determination relative to the degree that external factors, such as chance, fate, and luck, shape their lives, where larger values signify less internal control and more external control; and Rosenberg Self-Esteem measures the self-esteem of the individual based on a 10-part questionnaire in 1979. The Illicit Index is constructed based on the answers to 20 questions in the 1980 survey, where 17 are questions about "delinquency" and 3 are about interactions with the "police." The delinquency questions cover issues associated with damaging property, fighting at school, shoplifting, robbery, using force to obtain things, assault, threatening to assault somebody, drug use, dealing drugs, gambling, etc. The "police" questions involve being stopped by the police, charged with an illegal activity, or convicted, all for activities other than minor traffic offenses. Entrepreneur is based on the 2010 survey question, "Do you consider yourself to be an entrepreneur (where an entrepreneur is defined by the questioner as someone who launches a business enterprise, usually with considerable initiative and risk)?" We obtain the residuals from a regression of Entrepreneur on education AFQT, Rosenberg Self-Esteem, Rotter Locus of Control, the Illicit Index, and year of birth. As indicated, we standardize many of the variables to have a mean of 0 and a standard deviation of 1. Applied for Patent is similarly calculated based on the 2010 survey question, "Has anyone, including yourself, ever applied for a patent for work that you significantly contributed to? The Online Data Appendix provides detailed variable definitions and information on the construction of the data set.

youths than salaried workers, the incorporated engage in still more.[12]

12. All of these differences are statistically significant when using simple cross group *t*-tests. These findings are consistent with the observations of Steve Wozniak, the co-founder of Apple, who hacked telephone systems early in his career, "I think that misbehavior is very strongly correlated with and responsible for creative thought" (Kushner 2012). Our findings are also consistent with the work of Horvath and Zuckerman (1993), Zuckerman (1994), and Nicolaou et al. (2008), who argue that personality traits influence sorting into entrepreneurship,

Furthermore, after working for a couple of decades, the incorporated are more likely to describe themselves as "entrepreneurs" and more likely to have contributed to a patent. The variable Entrepreneur equals 1 if the individual responds "yes" to the question in the 2012 survey: "Do you consider yourself to be an entrepreneur (where an entrepreneur is defined by the questioner as someone who launches a business enterprise, usually with considerable initiative and risk)?" Since Entrepreneur is obtained decades after a person becomes prime age, we calculate the residuals from a regression of Entrepreneur on education, AFQT, Rosenberg Self-Esteem, Rotter Locus of Control, the Illicit Index, and year of birth. We then standardize these residuals to obtain Entrepreneur (residual standardized), which has a mean of 0 and a standard deviation of 1. We follow the same procedure to calculate Applied for Patent (residual standardized). As shown in Panel C of Table II, Entrepreneur (residual standardized) equals 1.2 for the incorporated and 0.69 for the unincorporated. The difference is even larger when examining patents. Applied for Patent (residual standardized) is 0.28 for the incorporated and only 0.03 for the unincorporated. These findings are consistent with our strategy of using the incorporated as a better proxy for those engaged in entrepreneurial activities than the aggregate group of self-employed.

II.D. *Does Incorporation Reflect Ex Post Sorting on Business Success?*

Using incorporation as an empirical proxy for entrepreneurship requires that the legal form of the business reflects the nature of the planned business activity, not simply the ex post success of the business. The concern is that businesses start as unincorporated firms and then incorporate if they are successful. To assess this concern, we exploit the long-term panel nature of the NLSY79 and examine self-employment spells. We define a self-employment spell as the full set of consecutive years in which a person is self-employed (either incorporated or unincorporated). For example, if a person is self-employed in 1991 and 1992, salaried in 1993, self-employed in 1994, and salaried in 1995, we define this person as having two self-employment spells. We examine all such spells in the NLSY79 sample, where some individuals experience more

---

and with Fairlie (2002), who shows that people who engaged in drug dealing as youths are more likely to become self-employed later in life.

TABLE III

SWITCHING BETWEEN UNINCORPORATED AND INCORPORATED SELF-EMPLOYMENT,
NLSY79

	Years			
Years as:	0	1	2	3 or more
Unincorporated before a business owner incorporates	84.5%	3.5%	5.0%	7.1%
Incorporated before a business owner unincorporates	98.1%	0.8%	0.5%	0.6%

Notes. This table provides information on the degree to which business owners switch the legal form of their businesses. For the statistics on the "years as unincorporated before a business owner incorporates," the following procedure is used: (i) consider self-employment spells in which a business owner ends the spell as incorporated self-employed, where a self-employment spell is one or more consecutive years in which an individual is self-employed (either incorporated or unincorporated) and (ii) compute for each spell the number of years the person was unincorporated before incorporating. The table reports the percentage of individuals for which the number of years is zero, one, two, or three or more. For example, the first column shows that for self-employment spells in which an individual ends the spell as an incorporated business owner, 84.5% started the spell as an incorporated business owner. An analogous procedure is followed for the statistics on "years as incorporated before a business owner unincorporates." As shown, for self-employment spells in which an individual ends the spell as an unincorporated business owner, 98.1% started the spell as an unincorporated business owner. Starting with the sample from Table I, the analyses in this table only include people who have a self-employment spell. See the Online Data Appendix for further details.

than one self-employment spell. If at the end of a self-employment spell the individual is an incorporated business owner, we determine how many years the individual was unincorporated self-employed before incorporating. Starting from the sample in Table I, Table III only includes people who have a self-employment spell. The results are virtually identical for the subsamples of males, whites, or white-males. Similarly, if at the conclusion of a self-employment spell the individual is an unincorporated business owner, we determine how many years the individual was incorporated before becoming unincorporated self-employed.

Table III shows that few people switch the legal forms of their businesses: people choose the legal form of the business when they choose to run it and rarely change afterward. In those cases when an individual ends a self-employment spell as an incorporated business owner, 85% of the time the person started the spell as an incorporated business owner. Most of the others switch in the first two years. Furthermore, 98% of those that end a self-employment spell as an unincorporated business owner also began the spell as unincorporated. These statistics are consistent with the view that individuals select the legal form of their businesses ex ante, not based on the ex post success of the endeavor. These analyses are also consistent with those in La Porta and Shleifer (2008, 2014), who find that few firms in developing economies transit from informal to formal.

II.E. Job Task Requirements

To assess whether the incorporated self-employed perform different tasks and run different types of businesses from the unincorporated, we use the U.S. Department of Labor's Dictionary of Occupational Titles (DOT) to measure the skills demanded of each occupation. The DOT was constructed in 1939 to help employment offices match job seekers with job openings. It provides information on the skills demanded of over 12,000 occupations. The DOT was updated in 1949, 1964, 1977, and 1991, and was replaced by the O*NET in 1998. Given the timing of our study, we use the 1991 DOT and confirm the results with the 1977 DOT.

The DOT aggregates information into several skill categories and we focus on three that are relevant for our study of entrepreneurship. For each category, it assigns a value between 0 and 10, where higher values signify that the job requires more of that skill. The first two skill categories measure the nonroutine cognitive skills demanded by particular jobs.

- Nonroutine Analytical indicates the degree to which the task demands analytical flexibility, creativity, reasoning, and generalized problem solving.
- Nonroutine Direction, Control, Planning indicates the degree to which the task demands complex interpersonal communications such as persuading, selling, and managing others.

The DOT also provides data on skills that align less directly with influential conceptions of entrepreneurship, notably Schumpeter (1911), Knight (1921), Lucas (1978), Baumol (1990), Murphy, Shleifer, and Vishny (1991), and Gennaioli et al. (2013).

- Nonroutine Manual measures the degree to which the task demands eye, hand, and foot coordination, which is high in such activities as landscaping, truck driving, carpentry, plumbing, and piloting an airline.

To link the DOT measures to the CPS and NLSY79 data, we follow Autor, Levy, and Murnane (2003) and use the codes provided on David Autor's website. We use the DOT to examine cross-sectional differences in the skill requirements of the incorporated and unincorporated and to measure differences in the types of businesses they run.

TABLE IV

JOB TASK REQUIREMENTS BY EMPLOYMENT TYPE

	All	Salaried	All	Self-employed Unincorporated	Incorporated
Panel A: CPS 1996–2012					
Job task requirements					
Nonroutine analytical	3.91	3.87	4.27	3.93	4.89
Nonroutine direction, control, planning	3.00	2.92	3.87	3.19	5.10
Nonroutine manual	0.99	0.99	0.98	1.08	0.80
Job task requirements last year (if salaried)					
Nonroutine analytical	4.04	4.03	4.18	3.82	4.68
Nonroutine direction, control, planning	3.15	3.14	3.50	2.82	4.45
Nonroutine manual	0.95	0.95	0.96	1.09	0.77
Panel B: NLSY79 1982–2012					
Job task requirements					
Nonroutine analytical	3.72	3.73	3.65	3.43	4.51
Nonroutine direction, control, planning	2.73	2.69	3.12	2.80	4.33
Nonroutine manual	1.05	1.03	1.19	1.25	0.95
Job task requirements last salaried job					
Nonroutine analytical	3.72	3.73	3.69	3.53	4.30
Nonroutine direction, control, planning	2.67	2.67	2.69	2.41	3.70
Nonroutine manual	1.05	1.03	1.17	1.23	0.97

Notes. The table presents summary statistics from the March Annual Demographic Survey files of the Census Bureau's CPS for the work years 1995 through 2012, for prime age workers (25 through 55 years old), and from the Bureau Labor of Statistics' National Longitudinal Survey of Youth 1979 (NLSY79) for workers who are least 25 years old between 1982 and 2012, as in Table I, for those with valid occupation codes. For Panels A and B, we use data on job task requirements from Autor, Levy, and Murnane (2003), who link data from the Dictionary of Occupational Titles with the occupational categories in the CPS. Nonroutine Analytical measures the degree to which the task demands analytical flexibility, creativity, and generalized problem solving, including tasks such as forming and testing hypotheses, making medical diagnoses, etc. Nonroutine Direction, Control, Planning measures the degree to which the task demands complex interpersonal communications such as persuading, selling, and managing others. Nonroutine Manual measures the degree to which the task demands eye, hand, and foot coordination, including landscaping, truck driving, carpentry, plumbing, and piloting a commercial airline. For "Job task requirements last year (if salaried)" in Panel A, we only include individuals who (a) are part of the two-year matched CPS panel and (b) were salaried workers in the previous year (230,330 observations). For "Job task requirements last salaried job" in Panel B, which is based on the NLSY79, we use information on a respondent's last salaried job (if any) (120,156 observations). See the Online Data Appendix for further details.

Table IV provides summary statistics of the job task requirements across employment types for the CPS and NLSY79. For the CPS, we do this for our main sample and for individuals who were salaried workers in the last survey. For the NLSY79, we do this for the main sample and when using information on a respondent's last salaried job (if any).

Table IV shows that (i) the incorporated engage in activities that demand greater nonroutine analytical skills (i.e., Nonroutine Analytical and Nonroutine Direction, Control, and Planning skills) than the unincorporated and salaried workers; (ii) the unincorporated engage in jobs that demand greater manual

skills (i.e., Nonroutine Manual skills) than the incorporated and salaried workers; and (iii) these sharp differences in the skills demanded of people who sort into incorporated and unincorporated self-employment exist before they become business owners. Aggregating the incorporated and unincorporated individuals blurs differences in their job task requirements.

<div align="center">

III. Differences between the Incorporated and Unincorporated

</div>

In this section, we assess whether people who become incorporated business owners (i) perform different activities from those who become unincorporated business owners and (ii) run different types of businesses.

We use multinomial logit regressions to assess whether people who perform jobs that demand a high-level of Nonroutine Analytical, Nonroutine Direction, Control, and Planning (DCP), or Nonroutine Manual skills are more likely to become incorporated business owners. We examine the sorting into employment types based on the job task requirements of the individual as a salaried worker in year $t - 1$ using the two-year matched panel of the CPS for work years 1995 through 2012, and further restrict the main sample to individuals who were salaried workers in $t - 1$ (as in Panel A of Table IV).

Specifically we estimate a multinomial logit model assuming that the log-odds of each worker conform to the following linear model:

$$(1) \qquad log \, \frac{P_{ijt}}{P_{ist}} = \alpha_j \, + \, \sum_{k \, = \, 1}^{3} \alpha_{NR,k,j} NR_{k,it-1} \, + \alpha_{X,j} X_{it-1}.$$

The dependent variable is the log-odds ratio of being an incorporated (unincorporated) business owner rather than a salaried worker, where P_{ijt} stands for the probability that person i is incorporated ($j = 1$) or unincorporated self-employed ($j = 2$) at time t, P_{ist} denotes the probability that the person is a salaried worker at time t, and k signifies the job task requirement category. $NR_{k,it-1}$ is a vector of three nonroutine job specific skill requirements (Analytical, DCP, and Manual) of person i's salaried job in year $t - 1$. X_{it-1} is a vector of regressors that includes demographics (race, gender), schooling, potential work experience (quartic), the

TABLE V

SELECTION INTO UNINCORPORATED AND INCORPORATED SELF-EMPLOYMENT

	(1) Unincorporated	(2) Incorporated
Job task requirements last year:		
Nonroutine analytical	−0.038**	0.055***
	(0.019)	(0.017)
Nonroutine direction, control, planning	−0.001	0.039***
	(0.006)	(0.008)
Nonroutine manual	0.037**	−0.139***
	(0.018)	(0.031)
Demographics:		
Years of schooling	0.011	0.055***
	(0.012)	(0.012)
Annual hours worked last year	−0.998***	0.418***
	(0.077)	(0.109)
Female	−0.366***	−0.734***
	(0.049)	(0.048)
Observations	230,330	230,330
Pseudo R-squared	0.99	0.99

Notes. This table reports multinomial logit estimates of the log-odds ratio of a salaried worker in $t − 1$, between the ages of 25 and 55, being unincorporated or incorporated self-employed rather than a salaried worker in year t. The sample excludes people who do not work either as salaried or self-employed, people with missing data on relevant demographics and labor market outcomes, and people living within group quarters. The analyses include the subsample of CPS observations, salaried or self-employed in year t, for which we have a matched, two-year panel over the work years 1995 through 2012 and restrict the sample to individuals who were salaried workers in the previous year ($t − 1$) as in Table IV, Panel A. Though unreported in the table, all specifications control for potential work experience (quartic), race, year, and state fixed effects. Data on job task requirements are from Autor, Levy, and Murnane (2003), who link data from the Dictionary of Occupational Titles with the occupational categories in the CPS. Nonroutine analytical measures the degree to which the task demands analytical flexibility, creativity, and generalized problem solving, including tasks such as forming and testing hypotheses, making medical diagnoses, etc. Nonroutine direction, control, planning measures the degree to which the task demands complex interpersonal communications such as persuading, selling, and managing others. Nonroutine manual measures the degree to which the task demands eye, hand, and foot coordination, including landscaping, truck driving, carpentry, plumbing, and piloting a commercial airline. Heteroskedasticity-robust standard errors clustered at the year-level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See the Online Data Appendix for further details.

number of hours worked in year $t − 1$, as well as state and year fixed effects.[13]

The estimates reported in Table V provide five messages about the sorting of people into incorporated and unincorporated self-employment based on the job task requirements of their

---

13. Potential work experience (pwe) equals age minus years of schooling minus 6 (or 0 if this computation is negative). The quartic includes pwe, pwe^2, pwe^3, and pwe^4.

previous jobs. First, people who open incorporated businesses were more likely to have been working in salaried jobs that demand greater nonroutine cognitive abilities than people who remained in salaried jobs. Second, the opposite is true of the unincorporated: people who open unincorporated businesses were less likely to have been working in salaried jobs demanding strong Nonroutine Analytical abilities than people who remained in salaried jobs. Third, people who open incorporated businesses were less likely to have been working in salaried jobs that required a high degree of Nonroutine Manual skills than people who remained in salaried jobs. Fourth, the results on the unincorporated are different: people who start unincorporated businesses tend to have worked in jobs requiring greater Nonroutine Manual skills than those that remained salaried workers. Fifth, the economic magnitudes are material. For example, consider a person who is working in a salaried job that demands about one-half of one standard deviation greater Nonroutine Analytical skills (0.94) than his counterpart. This gap is about the same as the difference between Nonroutine Analytical requirements of incorporated and unincorporated business owners in the year before they switch into self-employment, which is obtained from Panel A of Table IV (4.68−3.82). Holding other things equal, including the other job task requirements, the coefficient estimates in Table V suggest that relative to his counterpart, the odds of this person becoming an incorporated business owner next period are approximately 9% greater than becoming an unincorporated business owner ($1.09 = \frac{1.05}{0.96} = \frac{\exp(0.94*0.055)}{\exp(0.94*-0.038)}$).

Table V also demonstrates that other factors, beyond a person's job task requirements as a salaried worker, account for selection into different employment types. More educated people are more likely to become incorporated business owners, and women are less likely to become self-employed, especially incorporated business owners.[14] While individuals who worked more hours as salaried workers have a greater probability of becoming incorporated self-employed, the opposite is true for the unincorporated self-employed.[15]

14. Carr (1996), Budig (2006), and Özcan (2011) examine the sorting of men and women into incorporated and unincorporated self-employment.

15. The pseudo *R*-squared in Table V is high because the regressions include state and year fixed effects. The results are robust to excluding these fixed effects as shown in the Online Appendix Table V.A. Furthermore, as shown in the

Turning from the individual to the business, we construct and analyze measures of the job task requirements of each business. Specifically, using our main CPS sample, we compute the hours-weighted job task requirements of all workers in each industry over the work years 1995 through 2012 for each of three categories of skills: (i) Nonroutine Analytical skills, (ii) Nonroutine Direction, Control, and Planning skills, and (iii) Nonroutine Manual skills. In Table VI, we list the top five and bottom five industries of these three categories of the hours-weighted job task requirements of industries. The rankings seem intuitively plausible. Taxicab service, trucking service, and logging are the top five industries demanding high levels of manual skills from their workers, but they are the bottom five industries demanding nonroutine analytical skills from those same employees. In turn, engineering and architectural services demand high levels of analytical skills from workers, while the legal services and accounting industries do not require much in the way of nonroutine manual skills from their workers. Later, we will use these measures to assess the link between the nature of the person and the nature of the business.

IV. WHO BECOMES AN ENTREPRENEUR?

In this section, we examine the cognitive and noncognitive traits associated with the self-sorting of individuals into different employment types. We use the unique attributes of the NLSY79 data to examine how the traits of individuals before they enter the prime age labor market account for subsequent career choices. Above, we focused on the skills demanded by particular jobs and industries. We now focus on the pre-labor market "supply" of human capital traits.

IV.A. Selection into Employment Types on Cognitive and Noncognitive Traits

To further assess the association between pre–labor market measures of cognitive and noncognitive traits and subsequent employment choices, we estimate a multinomial logit model assuming that the log-odds of each response conform to the following

---

Online Appendix Table V.B, the Table V results on the demographic variables hold when including occupation fixed effects for the individual's job last year, rather than conditioning on the job task requirements of the occupation.

TABLE VI
TOP AND BOTTOM INDUSTRIES BY NONROUTINE JOB TASK REQUIREMENTS

Nonroutine analytical industry		Nonroutine direction, control, planning industry		Nonroutine manual industry	
Panel A: Top industries					
Engineering and architectural services	6.56	Educational services	6.23	Taxicab service	4.26
Accounting, auditing, and bookkeeping services	5.84	Engineering and architectural services	6.05	Trucking service	3.52
Miscellaneous professional and related services	5.34	Accounting, auditing, and bookkeeping services	5.05	Street railways and bus lines	3.14
Security and commodity brokerage and investment companies	5.16	Advertising	4.93	Logging	2.66
Aircraft and parts	5.08	Theaters and motion pictures	4.90	Water transportation	2.16
Panel B: Bottom industries					
Private households	0.99	Private households	0.45	Legal services	0.07
Logging	2.07	Taxicab service	0.57	Accounting, auditing, and bookkeeping services	0.09
Taxicab service	2.18	Postal service	0.98	Insurance	0.16
Trucking service	2.19	Trucking service	1.04	Security and commodity brokerage and investment companies	0.18
Laundering, cleaning, and dyeing services	2.30	Legal services	1.25	Banking and credit agencies	0.21

Notes. This table reports the top and the bottom five industries in each of three categories of job task requirements from the Dictionary of Occupational Titles. For each industry, we compute the hours-weighted job task requirements of people working in the industry over the work years 1995 through 2012 using the March Annual Demographic Survey files of the Census Bureau's CPS for the same sample as in Table I. We exclude industries with less than 1,000 observations. We do this for three categories of skills for each: (i) Nonroutine analytical measures the degree to which the task demands analytical flexibility, creativity, and generalized problem solving, including tasks such as forming and testing hypotheses, making medical diagnoses, etc.; (ii) Nonroutine direction, control, planning measures the degree to which the task demands complex interpersonal communications such as persuading, selling, and managing others; and (iii) Nonroutine manual measures the degree to which the task demands eye, hand, and foot coordination, including landscaping, truck driving, carpentry, plumbing, and piloting a commercial airline. See the Online Data Appendix for further details.

linear model:

$$log \frac{P_{ijt}}{P_{ist}} = \alpha_j + \alpha_{A,j} AFQT_i + \alpha_{I,j} Illicit_i + \alpha_{AI,j} AFQT_i \cdot Illicit_i$$

(2) $$\qquad\qquad + \alpha_{NC,j} NC_i + \alpha_{X,j} X_{it}.$$

The dependent variable is the log-odds ratio of being an incorporated (unincorporated) business owner rather than a salaried worker, where P_{ijt} is the probability that person i is incorporated ($j = 1$) or unincorporated self-employed ($j = 2$) at time t and P_{ist} denotes the probability that the person is a salaried worker. We focus on cognitive ability ($AFQT$) and noncognitive (NC) traits: the Rotter locus of control indicator, the Rosenberg self-esteem measure, and Illicit Activity. We also include an interaction between AFQT and Illicit. Therefore, throughout the remainder of our analyses, we exclude observations with missing values on the Illicit Activity Index from the main NLSY79 sample. All specifications control for gender, race, year-of-birth, and potential experience. When we introduce family traits, we also control for the Father's Education, Mother's Education, Family Income in 1979, and Two Parent Family (14).[16] By examining person-year observations, each person's "employment type" is defined by the number of years spent in each employment type. The errors are clustered at the individual level.

We report our findings in Table VII. In column (1), the logit model assesses the probability of self-employment versus salaried. All other columns report multinomial logit coefficient estimates. In columns (2)–(5), the comparison is between unincorporated self-employment and salaried, where column (5) repeats the analyses for column (4) using only the sample of whites; and in columns (6)–(9), the regressions provide estimates of the impact of each trait on the probability that the person is incorporated relative to being a salaried worker, where column (9) repeats the analyses for column (8) using only the sample of whites.

16. Though unreported in the table, when we control for family traits, we include a set of dummy variables for individuals with missing family income (for which we impute the average value in the sample) and missing parental education (for which we impute values based on the other parent's education and the average for the sample if no parental education is reported). The findings reported in Table VII are robust to the exclusion of observations with imputed family traits as shown in the Online Appendix Table VII.B.

TABLE VII

SELECTION INTO EMPLOYMENT TYPES ON COGNITIVE, NONCOGNITIVE, AND FAMILY TRAITS

Self-employment by type:	All (vs. salaried)	By type (vs. salaried)							
		Unincorporated				Incorporated			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cognitive and noncognitive traits									
AFQT	0.076	−0.046	−0.042	−0.105	−0.038	0.618***	0.576**	0.070	−0.121
	(0.115)	(0.124)	(0.124)	(0.132)	(0.163)	(0.235)	(0.237)	(0.261)	(0.294)
Illicit	0.078***	0.070**	0.133***	0.123**	0.160**	0.122**	−0.023	−0.045	−0.120
	(0.027)	(0.029)	(0.048)	(0.048)	(0.067)	(0.055)	(0.093)	(0.098)	(0.124)
Rosenberg score	0.031	−0.007	−0.009	−0.015	−0.006	0.211***	0.216***	0.190***	0.191***
	(0.029)	(0.031)	(0.031)	(0.031)	(0.040)	(0.059)	(0.059)	(0.060)	(0.069)
Rotter score	−0.097***	−0.089***	−0.087***	−0.086***	−0.085**	−0.141**	−0.144***	−0.130**	−0.135*
	(0.028)	(0.030)	(0.030)	(0.030)	(0.038)	(0.056)	(0.056)	(0.056)	(0.071)
AFQT*Illicit			−0.163	−0.150	−0.230*		0.306*	0.339**	0.461**
			(0.104)	(0.104)	(0.126)		(0.157)	(0.163)	(0.189)
Demographics									
Black	−0.560***	−0.504***	−0.501***	−0.538***		−0.887***	−0.898***	−0.784***	
	(0.072)	(0.075)	(0.075)	(0.077)		(0.164)	(0.165)	(0.168)	
Hispanic	−0.318***	−0.332***	−0.328***	−0.266***		−0.253	−0.260	0.042	
	(0.076)	(0.079)	(0.079)	(0.085)		(0.166)	(0.167)	(0.175)	
Female	−0.340***	−0.260***	−0.261***	−0.267***	−0.220***	−0.727***	−0.724***	−0.707***	−0.755***
	(0.055)	(0.059)	(0.059)	(0.059)	(0.075)	(0.119)	(0.119)	(0.119)	(0.146)

TABLE VII
(CONTINUED)

Self-employment by type:	All (vs. salaried)	By type (vs. salaried)							
		Unincorporated				Incorporated			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Family traits									
Family income in 1979				−0.080	−0.094			0.443***	0.437**
				(0.094)	(0.108)			(0.156)	(0.177)
Mother education				0.017	0.011			0.086***	0.117***
				(0.014)	(0.019)			(0.027)	(0.035)
Father education				0.010	0.020			0.008	−0.013
				(0.011)	(0.014)			(0.022)	(0.028)
Race/ethnicity	All	All	All	All	Whites	All	All	All	Whites
Observations	125,166	125,166	125,166	125,166	69,503	125,166	125,166	125,166	69,503
Pseudo R-squared	0.0276	0.0302	0.0306	0.0345	0.0268	0.0302	0.0306	0.0345	0.0268

Notes. This table reports multinomial logit estimates of the log-odds ratio of an individual being unincorporated or incorporated self-employed rather than a salaried worker. We use the sample in Table I and further exclude observations with missing values for Illicit. Though unreported in the table, all specifications control for year of birth and potential experience. When we introduce family traits, in columns (4), (5), (8), and (9), we also control for (but do not report the coefficient estimates on) (i) whether the respondent lived in a two-parent family at the age of 14, (ii) dummy variables for individuals with imputed family income, and (iii) dummy variables for imputed parental education, as defined in the Online Data Appendix. Reported standard errors (in parentheses) are corrected for heteroskedasticity and clustered by individual. The symbols ***, **, and * signify significance at the 1%, 5%, and 10% levels, respectively. See the Online Data Appendix for further details.

Several findings emerge. First, the incorporated self-employed tend to be white, male, people with high self-esteem, individuals with a strong sense of controlling one's future (i.e., a low Rotter locus of control score), individuals with high AFQT scores, those who engage in more illicit activities as youths, children of high-income parents, and people with well-educated mothers. The economic magnitudes are large. For example, holding other things constant, the odds of a woman becoming an incorporated business owner rather than a salaried employee are half of those of a similar male, which is obtained from the regression (8) estimates (i.e., $0.49 = \exp(-0.707)$). As another example, the odds of becoming incorporated self-employed rather than a salaried employee for a person with an AFQT score in the 60th percentile are 6.4% higher than for a person with the median AFQT score based on the regression (6) estimates.[17]

Second, family income predicts entrepreneurship. The coefficient estimates indicate that a $100,000 increase in family income—which is enough to boost somebody from the 10th to the 90th percentile—is associated with a more than 55% increase in the odds of becoming incorporated self-employed relative to those of becoming a salaried employee, after controlling for the person's cognitive and noncognitive traits, and other characteristics of the person's family environment. To the extent that one views family income as a proxy for credit constraints after controlling for other factors, these results indicate that difficulties in obtaining finance materially influence incorporated self-employment but not unincorporated self-employment.[18]

Third, people who have both high AFQT scores and high Illicit Activity Index values are much more likely to become incorporated business owners. For example, compare two people who are the same except for their AFQT and Illicit values. The first has the sample average value of Illicit (0) and the median value of AFQT (0.50), so that AFQT*Illicit equals 0. The second person, the "smart and illicit" person for this example, has one-quarter of one standard deviation above the mean value of Illicit (0.25) and is at the 75th percentile of the AFQT distribution (0.75), so that

17. AFQT was divided by 100 for the calculations in Table VII, so 1.0637= exp{0.618*(0.6–0.5)}.

18. For research on liquidity constraints and entrepreneurship, see, for example, the influential research by Evans and Jovanovic (1989), Holtz-Eakin, Joulfaian, and Rosen (1994), Blanchflower and Oswald (1998).

AFQT*Illicit is about 0.1875 (= 0.25*0.75). Then the odds of the smart and illicit person becoming an incorporated self-employed business owner rather than a salaried employee are 6.6% greater (exp {0.339*0.1875)}) than the first person. The mixture of high learning aptitude and disruptive, "break-the-rules" behavior is tightly linked with entrepreneurship.

Fourth, Table VII again emphasizes the differences in the prelabor market characteristics of people who become incorporated and unincorporated self-employed business owners. While the unincorporated also tend to engage in more illicit activities as youths than salaried workers, they do not have higher AFQT scores or self-esteem values; and, they do not come from particularly high-income or well-educated families. The combination of "smart" and "illicit" traits only boosts the probability of becoming incorporated self-employed.[19]

Fifth, Table VII also shows that using only the sample of whites yields very similar results. That is, among whites, the incorporated self-employed tend to be male, people with high self-esteem, low Rotter locus of control scores, have both high AFQT and Illicit Activity values, and come from high-income, highly educated families.

IV.B. Traits, Employment Types, and Job Task Requirements

We next examine the sorting of individuals into different types of business activities based on their early-determined traits and their choice of whether to establish an incorporated or

---

19. The NLSY79 data provide an opportunity to quantify the role of sorting on typically unobserved labor market skills since 90% of the people in our sample of full-time working adults are salaried workers at some point in their careers. Thus, we study the linkages between comparative success as a salaried worker and sorting into incorporated and unincorporated self-employment. Specifically, we first compute each individual's adjusted hourly wage as a full-time salaried employee by running a wage regression that controls for experience as well as year and individual effects and use the estimated individual effects as adjusted wages. We then run a new battery of multinomial logit regressions to assess whether adjusted wages and its interaction with Illicit explain sorting into employment types. As shown in the Online Appendix Table XIII, we discover that (i) there is negative sorting into the aggregate category of self-employment on adjusted wages, but this reflects positive sorting into incorporated self-employment and negative sorting into unincorporated self-employment and (ii) comparatively successful salaried workers who were also heavily engaged in disruptive activities as youths have higher propensities to become incorporated business owners later in life, that is, the interaction term enters positively and significantly.

unincorporated business. To do this, we match the cognitive and noncognitive traits of the business owner before he enters the labor market with the legal form and type of his business. We measure business type by the job task requirements of those employed by the business's industry, as previously defined when presenting Table VI.

Table VIII provides regressions in which the dependent variable is a measure of the job task requirements of the industry in which each individual works. The reported explanatory variables are dummy variables of whether the individual is an incorporated or unincorporated business owner, where salaried employment is the excluded group. The regressions also control for individual and year fixed effects, and a quartic in potential experience. We further restrict the sample to individuals who were salaried in the last NLSY79 survey, that is, in year $t - 2$ with valid industry codes. Thus, we compare people who remain salaried with those who switch into incorporated or unincorporated self-employment. We further restrict our NLSY79 analyses to white males since approximately 80% of the self-employed are white, and most of these are men (see Table I).[20]

To shed empirical light on whether individuals with particular combinations of cognitive and noncognitive traits establish different types of businesses when they start incorporated or unincorporated businesses, we examine four subsets of individuals. First, given the analyses above, we examine "smart and illicit" individuals, who have above the median values of both AFQT and Illicit (AFQT > 50 and Illicit > 0). Second, as a natural counterpart, we also examine individuals who are "not smart and illicit," that is, individuals with below (or equal to) the median values of either AFQT or Illicit (AFQT ≤ 50 or Illicit ≤ 0). Third, to get a sense of the intensive margin, we also examine the "very smart and illicit," individuals with above the 75th percentile AFQT scores and an Illicit index value greater than the median (AFQT > 75 and Illicit > 0). Finally, to assess whether it is just "smart," we examine the "very smart but not illicit", individuals who have above the 75th percentile AFQT scores but below (or equal to) the median values of the Illicit index (AFQT > 75 and Illicit ≤ 0).

We find that when "smart and illicit" individuals run incorporated businesses, they tend to be in industries that demand

---

20. The findings, however, are robust to including women and minorities, as shown in Online Appendix Table VIII.B.

TABLE VIII

DIFFERENCES IN JOB TASK REQUIREMENTS OF NEW BUSINESSES BY INDIVIDUAL TRAITS

The task requirements of the industry of the new business

	Nonroutine analytical industry				Nonroutine direction, control, planning industry				Nonroutine manual industry			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Incorporated	0.045	0.097	0.290***	0.078	−0.030	0.212**	0.541***	0.054	0.037	−0.033	−0.127*	−0.066
	(0.060)	(0.064)	(0.078)	(0.120)	(0.096)	(0.101)	(0.140)	(0.155)	(0.058)	(0.049)	(0.068)	(0.084)
Unincorporated	−0.060*	0.059	0.094	0.036	−0.246***	−0.126*	−0.217**	−0.283**	0.233***	0.090**	0.046	0.154***
	(0.032)	(0.055)	(0.091)	(0.085)	(0.045)	(0.075)	(0.110)	(0.124)	(0.028)	(0.045)	(0.074)	(0.059)
Sample												
AFQT	≤50	>50	>75	>75	≤50	>50	>75	>75	≤50	>50	>75	>75
Illicit index	or ≤0	and >0	and >0	and ≤0	or ≤0	and >0	and >0	and ≤0	or ≤0	and >0	and >0	and ≤0
Observations	22,542	6,870	3,504	5,483	22,542	6,870	3,504	5,483	22,542	6,870	3,504	5,483
R-squared	0.587	0.602	0.633	0.598	0.561	0.580	0.590	0.613	0.604	0.575	0.593	0.607

Notes. The dependent variable is the job task requirements of the industry in which the individual works, either as a business owner or as an employee. The industry job task requirements are computed from the CPS, as in Table VI. Starting from the Table I sample, these analyses only include prime age white male workers, who were salaried in year *t* − 2, and excludes observations with missing industry codes and missing values of Illicit. Incorporated and Unincorporated are indicators that equal 1 if the individual is incorporated self-employed or unincorporated self-employed respectively in year *t* and 0 otherwise. The regressions control for individual effects, a quartic in experience, and year fixed effects. Heteroskedasticity-robust standard errors clustered at the year-level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See the Online Data Appendix for further details.

comparatively high nonroutine cognitive skills from workers, and this tendency is even stronger for the "very smart and illicit." By comparing regressions (2) and (3) and (6) and (7), notice that the estimated coefficient on Incorporated is more than twice as large for the sample of individuals with AFQT > 75 and Illicit > 0 than it is for the sample of individuals with AFQT > 50 and Illicit > 0. Also, notice that the "very smart but not illicit" group of incorporated business owners does not have a stronger tendency to open high nonroutine cognitive businesses. The nature of the individual as a youth helps account for the type of incorporated business he runs later in life. Table VIII also provides information on the unincorporated. When "smart and illicit" individuals become unincorporated business owners, the businesses are not disproportionately in industries demanding strong analytical skills from workers. Rather, when most types of people open unincorporated businesses, they tend to be in industries that demand strong manual skills.

V. Do Entrepreneurs Earn More?

Using incorporated self-employment as a proxy for entrepreneurship, we examine whether (i) individuals earn more when they choose to become entrepreneurs than they were earning as salaried workers, (ii) the same human capital traits that account for who becomes an entrepreneur also explain success as an entrepreneur, (iii) the change in earnings associated with entrepreneurship exists across the full distribution of earnings, and (iv) the volatility of earnings rises relative to the increase in earnings when individuals become entrepreneurs.

In interpreting these analyses, it is worth emphasizing that entrepreneurship is not random and the change in earnings associated with becoming an entrepreneur may differ systematically with the human capital traits of the business owner. Indeed, the core findings presented above emphasize that people self-select into incorporated or unincorporated self-employment based on early-determined cognitive and noncognitive traits and the planned nature of the business. Therefore, in assessing whether entrepreneurs earn more, we do not aim to evaluate the causal impact on earnings of randomly assigning somebody to be an incorporated or unincorporated business owner. Rather, we aim to evaluate whether those who choose to become entrepreneurs earn more as business owners than they were earning as salaried workers.

To frame the analyses, consider the linear earnings equation:

$$(3) \qquad E_{it} = \beta_0 + \beta_I I_{it} + \beta_U U_{it} + \beta_X X_{it} + \varepsilon_{it},$$

where E_{it} equals the earnings of individual i at time t. To allow for nonpositive self-employment earnings, we examine earnings, not the log of earnings. I_{it} equals 1 if individual i is incorporated self-employed in period t and 0 otherwise, and U_{it} is a similarly defined dummy variable for when individual i is an unincorporated business owner. β_I and β_U are the gains in earnings associated with incorporated and unincorporated self-employment respectively relative to salaried earnings. X_{it} includes Mincerian characteristics and other relevant controls explained below. ε_{it} is the error term that can be decomposed into time-invariant individual fixed effects (θ_i) and time-varying individual influences ($a_i(t)$), along with a person-time shock to earnings (ϑ_{it}):

$$(4) \qquad \varepsilon_{it} = \theta_i + a_i(t) + \vartheta_{it}.$$

When excluding individual effects, the estimated β_I and β_U parameters provide unbiased measures of the differences in residual earnings for individuals who run incorporated or unincorporated businesses respectively relative to salaried employees. When including individual effects in the estimation of equation (3), the estimates for β_I and β_U yield unbiased estimates of the differences in residual earnings for individuals who choose to run incorporated and unincorporated businesses respectively relative to when they work as salaried employees.

V.A. Annual and Hourly Earnings

We now assess whether entrepreneurs earn more using data from the NLSY79.[21] The dependent variable is either annual or hourly earnings. We examine both annual and hourly earnings

---

21. All of the results hold when using the CPS, as shown in the Online Appendix Table IX and IX.B. Although the NLSY79 includes a smaller cross-section of individuals than the CPS, our earnings analysis focuses on the NLSY79 as it follows young adults over many years. We use the NLSY79's long panel to account for time-varying person-specific influences ($a_i(t)$) and to address potential biases from selection into and out of employment types, between and within employment spells, in estimating the expected gains relative to their earnings as salaried workers.

because the self-employed might have greater flexibility than others in adjusting work hours, as emphasized by Hurst and Pugsley (2011). We conduct the analyses both in levels and first differences and use both OLS and quantile (median) regressions. Since the NLSY79 survey is conducted every other year since 1994, the differencing is done between t and $t − 2$ for all years. We control for Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of cognitive and noncognitive traits (AFQT, Rosenberg self-esteem, Rotter locus of control, and the Illicit Activity Index), and year fixed effects.[22] When we control for individual fixed effects or take first differences, the time-invariant control variables drop from the analyses.[23] Throughout the earnings analyses, we restrict our main NLSY79 sample to white, male, full-time workers with nonmissing values on the Illicit Activity Index.[24]

The results in Table IX emphasize four messages. First, on average and at the median, the incorporated business owner earns more per annum and per hour than his salaried and unincorporated counterparts. For example, consider the median analyses conducted in levels and without conditioning on individual effects (columns (5) and (13)). The estimates indicate that the median incorporated business owner's annual residual earnings are $23,941 greater than those of the median salaried worker with the same observable characteristics and median residual hourly earnings are $5.32 greater. These estimates are large, when compared to the median earnings of salaried workers. As shown in the row labeled "% difference from salaried worker," our estimates indicate that the median residual annual earnings of the incorporated self-employed are 49% greater—and hourly earnings are 26% greater—than those of the median salaried worker. The corresponding OLS estimates (columns (1) and (9)) are much greater. For example, the estimated annual residual earnings of the

22. The six educational attainment categories: (i) high school dropouts: less than 12 years of schooling, (ii) GED degree, (iii) high school graduates: 12 years of schooling, (iv) had some college education: 13–15 years of schooling, (v) college education: 16 years of schooling, (vi) advanced studies: 17+ years of schooling. These are measured at the end of the respondent's educational experience, so that they do not vary over time for a respondent.

23. The polynomial expression for potential work experience in the first difference regressions is accordingly adjusted to be cubic.

24. In the levels analyses there are 23,657 observations, but the number of observations drops to 17,479 when using first differences.

TABLE IX
EARNINGS AND INDIVIDUAL EFFECTS

	OLS				Median			
	Levels		1st difference		Levels		1st difference	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Annual earnings								
Incorporated	45,926***	17,446***			23,941***	5,378***		
	(6,546)	(3,885)			(3,459)	(621)		
ΔIncorporated			12,592**	11,219**			3,953***	3,351***
			(5,748)	(5,201)			(519)	(364)
Unincorporated	8,893***	5,417***			−687	−367		
	(2,961)	(1,809)			(1,042)	(478)		
ΔUnincorporated			2,580	2,563			−728*	−399
			(2,183)	(2,473)			(393)	(327)
% Difference from salaried worker								
Incorporated	75%	29%	21%	18%	49%	11%	8%	7%
Unincorporated	15%	9%	4%	4%	−1%	−1%	−1%	−1%
Individual fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	23,657	23,657	17,479	17,479	23,657	23,657	17,479	17,479
R-squared	0.253	0.631	0.011	0.082	0.132	0.111	0.016	0.010

TABLE IX
(CONTINUED)

| | OLS | | | | Median | | | |
| | Levels | | 1st Difference | | Levels | | 1st Difference | |
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Panel B: Hourly earnings								
Incorporated	13.141***	4.384***			5.317***	0.978***		
	(2.250)	(1.407)			(1.094)	(0.311)		
ΔIncorporated			4.168**	3.846**			1.350***	0.928***
			(1.975)	(1.825)			(0.225)	(0.222)
Unincorporated	0.376	0.739			−2.737***	−0.849***		
	(1.062)	(0.665)			(0.356)	(0.224)		
ΔUnincorporated			0.013	0.004			−0.739***	−0.554***
			(0.808)	(0.893)			(0.185)	(0.175)
% Difference from salaried worker								
Incorporated	52%	18%	17%	15%	26%	5%	7%	5%
Unincorporated	2%	3%	0%	0%	−13%	−4%	−4%	−3%
Individual fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	23,657	23,657	17,479	17,479	23,657	23,657	17,479	17,479
R-squared	0.255	0.625	0.011	0.078	0.136	0.110	0.012	0.006

Notes. This table reports OLS and median regression results of both annual earnings and hourly earnings on employment type, using data from the NLSY79 for years 1982 through 2012 for the sample of white male, full-time workers between the ages of 25 and 55. We exclude observations with missing demographics (gender, race, education and potential experience) or missing values for AFQT, Rosenberg Self-Esteem, Rotter locus of control, and Illicit. The dependent variable in the level specifications is annual (hourly) earnings. Since the NLSY79 survey is conducted every other year since 1994, the differencing in the 1st difference specifications is done between t and $t − 2$ for all years. The dependent variable in the first difference specification is the change in annual (hourly) earnings between t and $t − 2$. ΔIncorporated and ΔUnincorporated equal the change in incorporated and unincorporated self-employment status respectively between t and $t − 2$. We control for Mincerian characteristics (a quartic expression for potential work experience and dummy variables for six education categories), measures of cognitive and noncognitive traits (AFQT, Rosenberg self-esteem, Rotter locus of control, and the Illicit Activity Index), and year fixed effects. When we control for individual effects or take first differences, the time-invariant control variables drop from the analyses. The polynomial expression for potential experience in the first difference specifications is accordingly adjusted to be cubic. When examining % differences from salaried workers, the statistics are based on the means for salaried workers in the OLS regressions and the medians for salaried workers in the median regressions. Heteroskedasticity robust standard errors clustered at the year-level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See the Online Data Appendix for details.

self-employed are $45,926 larger than those of the average salaried worker.

Second, after controlling for individual fixed effects, people who become incorporated business owners, on average and at the median, experience a material increase in annual and hourly earnings. For example, regressions (2) and (10) indicate that mean residual annual earnings of an individual increase by 29% when he becomes an incorporated business owner and his hourly earnings rise by 18%, relative to the earnings of an average salaried worker. Thus, he earns more per hour, works more hours, and earns much more per annum after switching into incorporated self-employment. The results also hold when examining median earnings (columns (6) and (14)), though the estimated relationships at the median are about one-third of the mean estimates, emphasizing the skewed distribution of incorporated self-employed earnings relative to salaried employment.[25]

Third, there is positive selection into incorporated self-employment on salaried earnings. To see this, compare the OLS regression estimates in column (1), which do not include individual fixed effects, with those in column (2), which condition on individual effects. The estimates suggest that on average, the person who runs an incorporated business earned $28,480 more per annum as a salaried worker than a person with the same observable characteristics who did not become an incorporated business owner. This gap is large, as it equals 46% of the earnings of an average salaried worker. Furthermore, this result holds for median hourly earnings (columns (13) and (14)).

Fourth, the results on the unincorporated self-employed are distinct. The median residual hourly earnings of somebody who switches from salaried work to unincorporated self-employment tends to fall by $0.85 (column (14)) relative to his hourly earnings as a salaried worker. However, the individual tends to work more hours, so his median residual annual earnings do not fall significantly (column (6)). Even though these analyses are limited

25. When we do not impose symmetry and allow the absolute value of the estimated change in earnings when somebody switches from salaried work to incorporated self-employment to differ from the estimated change when somebody switches from incorporated self-employment to salaried work, all of the reported results hold. Indeed, we find that when somebody returns to salaried work after incorporated self-employment, he returns to a wage rate that is very similar (in real terms) to the salary he had before becoming self-employed.

to full-time, full-year workers, the self-employed still work more hours than they were working as salaried employees. Furthermore, there is evidence that the median unincorporated business owner is negatively selected on his salaried earnings into self-employment. For example, the estimated drop in median hourly earnings from unincorporated self-employment is much lower in absolute terms when accounting for individual fixed effects, as shown in columns (13) and (14).[26]

The NLSY79 allows us to control for, and assess the importance of, time-varying person-specific influences $(a_i(t))$. The concern is that people might select into incorporated self-employment based on trends in their earnings. If people with steeper earnings profiles have a greater propensity to become incorporated business owners, then some of the estimated increase in earnings associated with switching from salaried work to incorporated self-employment could reflect this trend rather than a change in earnings associated with entrepreneurship. To assess this possibility, we control for both individual-specific factors (by conducting the analyses in first differences) and an individual-specific linear time trend (by including an individual fixed effect in the first difference regression).

The results hold, with only minor changes in the estimated parameters, when controlling for person-specific linear trends. This can be seen by comparing regressions (3) and (4), (7) and (8), (11) and (12), and (15) and (16) in Table IX. When examining first-differences in earnings while controlling for individual fixed effects, we continue to find that the earnings of an individual tend to rise—relative to the individual's trend line—when switching from salaried work to incorporated self-employment as reported in columns (4), (8), (12), and (16). Again, the estimated effects are slightly smaller for hourly earnings, indicating that people tend to work more hours when they become self-employed, even among this sample of full-time workers.[27]

26. As demonstrated below when we examine the full distribution of earnings, the positive selection into incorporated self-employment holds across virtually all percentiles, but selection into unincorporated is more nuanced.

27. These results on earnings are robust to four additional concerns. First, we were concerned that something odd could be happening during the year of incorporation. Thus, we omitted the two years before and the two years after incorporation and confirm the findings. Second, we were concerned that individuals buying into businesses in which they were working as salaried workers, rather than starting their own business, drove the results. This is not the case. Virtually all of the

V.B. *Selection into and out of Self-Employment*

We next use self-employment spells to assess the degree to which the estimated positive relationship between entrepreneurship and earnings is influenced by several selection concerns. One concern is positive selection into incorporated self-employment on earnings as an unincorporated business owner. Perhaps businesses start as unincorporated enterprises and then the successful ones incorporate. Although we showed in Table III that few people switch the legal form of their businesses within a self-employment spell, these few switchers could influence the estimated relationship between earnings and incorporated self-employment.

To account for selection into incorporated self-employment, we incorporate information on employment spells and rewrite the earnings equation (3) as follows:

$$(5) \qquad E_{its} = \beta_0 + \beta_I I_{its} + \beta_U U_{its} + \beta_X X_{its} + \varepsilon_{its}.$$

E_{its} equals the earnings of individual i at time t in employment spell s. An employment spell is the full set of consecutive years as either a salaried or self-employed worker. Since we only consider full-time workers, individuals are either salaried or self-employed in each period. Thus, if somebody is always salaried, the person will experience just one employment spell. If a person is salaried for 10 years, self-employed for 1, and then salaried for 5 more years, the person experiences three employment spells. I_{its} equals 1 if individual i in spell s during period t is incorporated self-employed and 0 otherwise. Thus, $I_{its} = 0$ in all salaried

---

switches into incorporation involve a change of firms. When we limit incorporation to situations in which a person changes firms, we get virtually identical results. Third, we were concerned that earnings growth might predict changes in employment type. Consequently, we examined the relationship between the change in hourly earnings between period $t - 2$ and $t-4$ and the change in employment type from period t to $t - 2$. If the change in earnings is associated only with a contemporaneous change in employment type, then we expect this regression to yield an insignificant coefficient. If, however, increases in earnings tend to precede transitions into incorporated, then we would expect to find a positive coefficient. There is not a statistically significant relationship between a change in earnings and subsequent shifts into incorporated self-employment. While earlier results document the positive sorting into entrepreneurship on earnings, the evidence does not indicate that jumps in earnings are good predictors of subsequent shifts into incorporation; rather, earnings jump when people become incorporated business owners. Fourth, as shown in the Online Appendix Table IX.D, the results hold when controlling for family traits.

employment spells. U_{its} is a similarly defined dummy variable for when individual i, in period t, and spell s is an unincorporated business owner. The variables in equation (5) are the same as those in equation (3). Equation (5) simply acknowledges formally that some people change between incorporated and unincorporated self-employment during a single self-employment spell. The error term (ε_{its}) can be decomposed into time-invariant and time-varying person-specific effects (θ_i and $a_i(t)$), a person-spell specific shock to earnings (η_{is}) and a zero-mean person-time shock to earnings (ϑ_{its}):

$$(6) \qquad \varepsilon_{its} = \theta_i + a_i(t) + \eta_{is} + \vartheta_{its}.$$

To assess whether positive selection into incorporated self-employment influences the estimated relationship between earnings and entrepreneurship, we eliminate within self-employment spell variation. We do this by defining a self-employment spell as incorporated or unincorporated based on the legal form of the business in the first year of the spell, so that

$$(7) \qquad E_{its} = \beta_0 + \beta_I I_{is} + \beta_U U_{is} + \beta_X X_{its} + \epsilon_{its},$$

where I_{is} equals 1 for all years of individual i's self-employment spell s if the individual started the spell as incorporated self-employed and 0 otherwise. Note that $I_{is} = 1$ for all years of the self-employment spell s when the individual starts the spell as incorporated self-employed regardless of whether he switches to unincorporated self-employment later in the spell. In Table X, we refer to I_{is} as "Spell starts as incorporated." Similarly, $U_{is} = 1$ for all years of individual i's self-employment spell s if the individual started the spell as an unincorporated self-employed business owner regardless of whether he switches to incorporated self-employment later in the spell. For individuals who do not switch from unincorporated to incorporated self-employment within an employment spell, $I_{its} = I_{is}$ for all t in the spell. However, for individuals who start a self-employment spell as an unincorporated business owner $(I_{is} = 0)$ and then incorporate later, $I_{its} - I_{is} = 1$ for some periods within spell s.

As shown in Table X, there is a large increase in earnings when an individual becomes an incorporated self-employed business owner relative to his earnings as a salaried employee even after accounting for positive selection into incorporated business

TABLE X
EARNINGS AND SELECTION IN AND OUT OF SELF-EMPLOYMENT

	OLS				Median			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Annual earnings								
Incorporated	17,446***				5,378***			
	(3,885)				(621)			
Unincorporated	5,417***				−367			
	(1,809)				(478)			
Spell starts as incorporated		14,064***	12,309***			4,604***	4,296***	
		(4,112)	(3,271)			(691)	(731)	
Spell starts as unincorporated		7,194***	4,995***			−78	−581*	
		(1,831)	(1,620)			(464)	(340)	
1st year of 1st spell incorporated				11,186***				4,334***
				(2,961)				(399)
1st year of 1st spell unincorporated				1,846				−1,467***
				(1,655)				(356)
% Difference from salaried worker								
Incorporated	29%	23%	20%	18%	11%	9%	9%	9%
Unincorporated	9%	12%	8%	3%	−1%	0%	−1%	−3%
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighted by inverse of years in spell	No	No	Yes	Yes	No	No	Yes	Yes
Observations	23,657	23,657	23,657	23,657	23,657	23,657	23,657	23,657
Spells	3,553	3,553	3,553	3,553	3,553	3,553	3,553	3,553
R-squared (Pseudo R2)	0.63	0.63	0.63	0.63	0.111	0.111	0.080	0.080

TABLE X
(CONTINUED)

	OLS					MEDIAN		
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Panel B: Hourly earnings								
Incorporated	4.384***				0.978***			
	(1.407)				(0.311)			
Unincorporated	0.739				−0.849***			
	(0.665)				(0.224)			
Spell starts as incorporated		3.520**	3.655***			0.925***	0.535**	
		(1.527)	(1.194)			(0.335)	(0.216)	
Spell starts as unincorporated		1.213*	0.691			−0.844***	−0.708***	
		(0.670)	(0.611)			(0.225)	(0.182)	
1st year of 1st spell incorporated				3.612***				0.667***
				(1.072)				(0.233)
1st year of 1st spell unincorporated				−0.300				−0.858***
				(0.623)				(0.143)
% Difference from salaried worker								
Incorporated	18%	14%	15%	14%	5%	5%	3%	3%
Unincorporated	3%	5%	3%	−1%	−4%	−4%	−3%	−4%
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighted by inverse of years in spell	No	No	Yes	Yes	No	No	Yes	Yes
Observations	23,657	23,657	23,657	23,657	23,657	23,657	23,657	23,657
Spells	3,553	3,553	3,553	3,553	3,553	3,553	3,553	3,553
R-squared (Pseudo R2)	0.625	0.625	0.623	0.623	0.110	0.110	0.077	0.077

Notes. This table reports OLS and median regression results of both annual earnings and hourly earnings on employment type, using data from the NLSY79 for years 1982 through 2012 for the sample of white male, full-time workers between the ages of 25 and 55, using the same core sample and variable definitions as in Table IX. Panel A provides results on annual earnings, while Panel B examines hourly earnings. "Spell starts incorporated" equals 1 for each year of a self-employment spell if the person starts the self-employment spell as an incorporated business owner and 0 otherwise. "Spell starts unincorporated" is defined analogously. "1st year of 1st spell incorporated" equals 1 for each year of all of the years an individual is self-employed if in the first year of the first employment spell the individual is incorporated self-employed and 0 otherwise. "1st year of 1st spell unincorporated" is defined analogously. Note that (i) an employment spell is the full set of consecutive years as either a salaried or self-employed worker (where individuals are either salaried or self-employed in each period since we only consider full-time workers) and (ii) a self-employment spell is the full set of consecutive years in which a person is self-employed (either incorporated or unincorporated). In the indicated specifications, the observations are weighted by the inverse of the number of years in the employment spell to give equal weight to each spell regardless of its length. All specifications control for individual effects, year effects, and a quartic expression for potential work experience. When examining % differences from salaried workers, the statistics are based on the means for salaried workers for the OLS regressions and the medians for salaried workers in the median regressions. Heteroskedasticity-robust standard errors clustered at the year-level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively. See the Online Data Appendix for further details.

ownership. The regressions control for individual effects, year effects, and potential work experience (quartic). Regression (1) in Table X replicates the findings from regression (2) of Table IV and shows that the estimated increase in earnings from an individual switching from salaried work into incorporated self-employment is \$17,446. In regression (2), we examine "Spell starts as incorporated" (I_{is}) and find that the estimated increase in earnings from an individual switching into a self-employment spell, in which the individual is incorporated self-employed in the first year, is \$14,064. This suggests that some of the estimated increase in earnings associated with becoming incorporated self-employed is that successful unincorporated businesses incorporate during the self-employment spell. But even when eliminating this positive selection, there is still a material boost in annual income of 23% relative to the average salaried worker in the sample when somebody switches from salaried employment to run an incorporated business. At the median, the point estimates are almost identical when accounting for selection into incorporated self-employment, indicating that only a few individuals make it big as unincorporated business owners and then switch to incorporated self-employment.

Another possible challenge to assessing whether entrepreneurs earn more is selection out of self-employment when the business is unsuccessful. As emphasized by Manso (2016), such "survivorship bias" would bias upward the estimated relationship between earnings and self-employment by giving more weight (in the form of systemically more observations) to successful self-employment spells than unsuccessful ones. Using I_{is} and U_{is} in equation (7) addresses selection across self-employment types but not selection out of self-employment.

To evaluate the empirical importance of selection out of self-employment, we weight the observations in the earnings regressions by the inverse of the number of observations in each employment spell, so that each employment spell gets equal weight. An incorporated, unincorporated, or salaried employment spell includes the full set of consecutive observations of that employment type. We report these results in regressions (3), (7), (11), and (15) of Table X, where we continue to (i) include individual fixed effects and (ii) use I_{is} and U_{is} to control for selection into incorporated and unincorporated self-employment.

All of the results hold after controlling for survivorship bias. Although the self-employed exercise the option to dropout and

return to salaried work when the business does not succeed, we find that the main findings hold after accounting for this type of selection. To see this compare the coefficient estimates between columns (2) and (3) for the OLS estimates and columns (6) and (7) for the median results. The point estimates are slightly lower and statistically indistinguishable.[28]

Another concern relates to selection into incorporated and unincorporated self-employment between spells. Since half of those who become self-employed have two or more self-employment spells, we were concerned that individuals choose to incorporate when they identify a particularly promising business opportunity and start unincorporated businesses when that is not the case. Under these conditions, the increase in earnings associated with incorporated self-employment might primarily reflect selection into incorporated self-employment on expected earnings and not a boost in earnings associated with the nature of the business or the human capital skills of the owner.

But this concern does not materialize in the data. First, we find that 84% of those individuals who have two or more self-employment spells choose to be either incorporated or unincorporated in all of those spells. There is very little variation in the legal form of businesses across an individual's self-employment spells. Second, as shown in Table X, we find no evidence that the few people who switch between incorporated and unincorporated self-employment across their different self-employment spells change our findings. We categorize all of an individual's self-employment observations by the first year of his first self-employment spell and redo the analyses, such that "1st year of 1st spell incorporated" equals 1 for each year of all of the years that an individual is self-employed, if in the first year of his first employment spell the individual is incorporated self-employed, and 0 otherwise. We define "1st year of 1st spell unincorporated" analogously. We then reestimate the earnings regressions and provide the results in

28. The panel nature of the NLSY79 data also provides an opportunity to provide greater insights on the earnings profiles of individuals who try self-employment and then return to salaried work. First, we discover that individuals who experiment with entrepreneurship and then return to salaried employment on average return to higher paying salaried jobs (hourly earnings) than they had before becoming incorporated business owners. Second, the results on unincorporated self-employment are different. After an individual becomes an unincorporated self-employed business owner, his future hourly earnings fall regardless of whether he returns to salaried employment.

columns (4), (8), (12), and (16). We continue to find that an individual's earnings rise appreciably when he becomes incorporated self-employed—even after removing the potential effects of individuals choosing different legal forms for their businesses across self-employment spells. These findings are unsurprising given our earlier results that (i) early-determined human capital traits account for the self-sorting of people into incorporated or unincorporated self-employment and (ii) those who become incorporated business owners engage in different types of activities and run different types of businesses from those who become unincorporated business owners.

V.C. Which Entrepreneurs Earn More?

We now use the earlier analyses on who becomes an entrepreneur to examine whether the same traits associated with selection into entrepreneurship are also associated with larger increases in earnings when an individual becomes an entrepreneur. To do this, we differentiate individuals by whether they are "smart and illicit," that is, whether they have both high AFQT scores and strong tendencies to break the rules as youths (AFQT > 50 and Illicit > 0), or whether they do not (AFQT $\leqslant 50$ or Illicit $\leqslant 0$).[29] Specifically, we take the first difference of equation (7) and conduct the analyses while separately examining the samples of smart and illicit individuals and all others:

$$(8) \qquad \Delta E_{its} = \delta_0 + \delta_1 \Delta I_{is} + \delta_2 \Delta U_{is} + \delta_3 \Delta X_{its} + u_{its}.$$

The change in individual i's incorporated self-employment status is ΔI_{is} and the change in individual i's unincorporated self-employment status is ΔU_{is}, where incorporated and unincorporated self-employment status are defined by the first year of the self-employment spell. As above, these first difference regressions include potential experience and year fixed effects.

Table XI provides estimates of the change in median residual annual and hourly earnings associated with switching into or out

29. We differentiate between the "smart and illicit" and others. When using the same demarcation employed in Table VIII, we find that the change in earnings associated with becoming an incorporated self-employed business owner is especially pronounced among the "very smart and illicit" and nonexistent among the "very smart but not illicit."

TABLE XI

THE CHANGE IN MEDIAN EARNINGS DIFFERENTIATING BY "SMART AND ILLICIT"

	(1)	(2)	(3)	(4)
△Incorporated	716**	6,996***		
	(285)	(547)		
△Unincorporated	105	−1,895**		
	(426)	(807)		
△Self-employed × nonroutine cognitive industry			−3,234***	8,163***
			(955)	(1,052)
△Self-employed			1,155**	−1,550*
			(452)	(937)
Nonroutine cognitive industry			1,139***	1,563***
			(193)	(361)
% Difference from salaried worker				
Incorporated	1%	12%		
Unincorporated	0%	−3%		
Self-employed in nonroutine industry			−4%	11%
Self-employed			2%	−3%
Sample	AFQT ⩽ 50 or Illicit ⩽ 0	AFQT > 50 & Illicit > 0	AFQT ⩽ 50 or Illicit ⩽ 0	AFQT > 50 & Illicit > 0
Pseudo R-squared	0.016	0.017	0.017	0.018
Observations	13,269	4,210	13,269	4,210

Notes. This table reports median regressions of the change in annual earnings on the change in employment type for white males working full-time, using data from the NLSY79 for years 1982 through 2012, the same sample as in Table IX. In specifications (1) versus (2) and (3) versus (4), the sample is split between individuals who have (a) AFQT ⩽ 50 or Illicit ⩽ 0 and (b) the smart and illicit with AFQT > 50 and Illicit > 0. In regressions (1) and (2), the main explanatory variables are the change in the incorporated and the unincorporated status over the past two years, where incorporated and unincorporated employment status are defined by the first year of the self-employment spell. A self-employment spell is the full set of consecutive years in which a person is self-employed (either incorporated or unincorporated). In regressions (3) and (4), the main explanatory variables are (i) the change in self-employment status interacted with a dummy variable of whether the business is in a Nonroutine Cognitive Industry or not, (ii) the change in self-employment status, and (iii) a dummy variable of whether the person works in a Nonroutine industry or not. A Nonroutine Cognitive industry is an industry that demands both above average values of Nonroutine Analytical skills (analytical flexibility, creativity, reasoning, and generalized problem solving) and Nonroutine Direction, Control, Planning skills (complex interpersonal communications such as persuading, selling, and managing others) from its workers. The change in self-employment status equals 1 if the person switches from salaried work in $t − 2$ to self-employment in t. The statistics for % difference from salaried workers are calculated for the corresponding group of salaried workers, for example, in specification (2), the change in annual earnings is computed relative to the median among salaried workers with AFQT > 50 and Illicit > 0, and in specification (4), the computations are done relative to the median among salaried workers with AFQT > 50 and Illicit > 0 in Nonroutine Cognitive Industries. All specifications control for a cubic expression in potential work experience and year fixed effects. Heteroskedasticity-robust standard errors clustered at the year-level are in parentheses, where *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. See the Online Data Appendix for further details.

of incorporated and unincorporated self-employment, where we differentiate by whether individuals are "smart and illicit" or not. To control for positive selection into incorporated self-employment from unincorporated self-employment, we continue to define a person's employment type by the first year of the employment spell. By conducting the earnings analyses in first differences and

defining self-employment type by the first year of the spell, the analyses control for both types of selection.

Table XI indicates that "smart and illicit" individuals who become incorporated business owners enjoy much larger increases in annual and hourly earnings than (i) individuals who do not have these particular combinations of traits and become incorporated business owners and (ii) smart and illicit individuals who become unincorporated business owners. That is, the same traits associated with selection into incorporated self-employment also account for the magnitude of the increase in earnings when an individual becomes an entrepreneur. For example, while the smart and illicit enjoy an almost $7,000 increase in median residual annual earnings when becoming incorporated business owners relative to their earnings as salaried workers (column (2)), others experience only a $716 increase (column (1)). The changes associated with the smart and illicit becoming incorporated self-employed are economically large. For example, the $7,000 increase in median residual earnings associated with a smart and illicit individual becoming an incorporated business owner is 12% of the median residual earnings of their salaried smart and illicit counterparts. The smart and illicit experience much bigger increases in earnings when they become incorporated business owners, in absolute and relative terms, than people with different traits.

The results on the unincorporated self-employed are very different and emphasize (i) the sharp distinction between entrepreneurship and other self-employment activities and (ii) the degree to which different combinations of traits are differentially valuable in different activities. In contrast to the findings on those who become incorporated business owners, Table XI indicates that smart and illicit individuals who become unincorporated self-employed experience a larger drop in hourly earnings than individuals with different traits who become unincorporated business owners. The combination of smart and illicit traits is positively associated with success as an entrepreneur but negatively associated with success in other self-employment activities.[30]

30. To the extent that underreporting of income is correlated with smart and illicit traits differentially among the incorporated and unincorporated self-employed, such that the underreporting gap between unincorporated and incorporated is larger among the smart and illicit, this could bias the results toward

There is a potential concern that smart and illicit individuals are more likely than others to start incorporated businesses when they expect earnings to be especially high and not simply when they start an entrepreneurial business. To assess whether the smart and illicit experience larger increases in income when they become entrepreneurs than individuals with other traits who open such businesses or than when the smart and illicit open nonentrepreneurial businesses, we assess whether the results hold without conditioning on the legal form of the business.

Thus, we evaluate whether smart and illicit people experience especially large boosts in earnings when they become self-employed in industries that demand high levels of nonroutine cognitive skills from workers. We define "Nonroutine cognitive industries" as those that demand above average values of both Nonroutine Analytical and Nonroutine Direction, Control, Planning skills from workers. The focus on nonroutine cognitive industries reflects our earlier argument that nonroutine cognitive activities are closely aligned with core conceptions of entrepreneurship, while strong manual skills are not. A key shortcoming with using this industry-level variation to assess gains in earnings associated with entrepreneurship is that the extra earnings from becoming self-employed in a nonroutine cognitive industry might reflect an industry effect rather than an "entrepreneurship" effect. By comparing smart and illicit individuals to others within and between industries our "difference-in-differences" external validity setting allows us to account for both the type of person and the type of industry effects on the change in earnings.

In regressions (3) and (4) of Table XI, the dependent variable remains the change in median annual earnings. Rather than including ΔI_{is} and ΔU_{is} as regressors, we use (1) ΔSelf-Employed $\times$ Nonroutine Cognitive Industry, which is the interaction between the change in the individual's self-employment status and whether the business is in a Nonroutine Cognitive Industry or not; (2) ΔSelf-Employed, which is the change in the individual's self-employment status; and (3) Nonroutine Cognitive Industry, which equals 1 if the person works, either as a salaried worker or self-employed business owner, in a Nonroutine Cognitive Industry. We provide estimates for two samples (i) people who

---

finding that smart and illicit traits yield bigger underreporting rewards when the person is an unincorporated business owner.

are not both smart and illicit and (ii) people who are both smart and illicit.

The results indicate that smart and illicit individuals who become self-employed business owners in Nonroutine Cognitive Industries tend to experience large increases in annual earnings, but individuals without those traits actually tend to experience a drop in earnings when they become self-employed in such industries. That is, the same smart and illicit traits that are positively associated with (i) selection into incorporated self-employment, (ii) selection into business ownership in nonroutine cognitive industries, and (iii) increases in earnings when a person becomes an incorporated business owner are also positively associated with the increase in earnings associated with individuals becoming self-employed in Nonroutine Cognitive Industries.

Indeed, as shown in regression (4), smart and illicit individuals earn more as self-employed only when they open businesses in nonroutine cognitive industries. The median gain in earnings from self-employment is approximately \$6,613 (\$8,163 − \$1,550), which is remarkably similar to those from incorporated self-employment. This cannot be attributed to a common industry self-employment effect (as the coefficient estimate on ΔSelf-Employed $\times$ Nonroutine Cognitive Industry is actually negative). While selection into self-employment in different industries is neither random nor exogenous to person traits, it is not contaminated by any ex ante or ex post selection into or out of incorporated vis-à-vis unincorporated self-employment. This robustness test is consistent with the view that a particular mixture of smart and illicit traits matters for success as an entrepreneur. The similarities between columns (2) and (4) reflect the finding that "smart and illicit" business owners are much more likely than others to choose the incorporated legal form, especially when they open businesses in Nonroutine Cognitive Industries.[31]

These findings on who succeeds as an entrepreneur contribute to existing research. Researchers examine the connection between the propensity for an individual to become self-employed and self-esteem, optimism, and a taste for novelty, as in Horvath and Zuckerman (1993), Zuckerman (1994), Nicolaou et al. (2008), and Hartog, Praag, and Sluis (2010). Lazear (2004, 2005) stresses

---

31. When smart and illicit people open businesses in Nonroutine Cognitive industries, they are twice as likely to choose the incorporated legal form than when other people open businesses in such industries (or in other industries).

that entrepreneurs must be "jacks-of-all-trades" to coordinate factor inputs successfully. Our work demonstrates that a special mixture of cognitive and noncognitive skills—the combination of outstanding abilities and disruptive tendencies—is strongly associated with entrepreneurial success.

V.D. Risk and the Dispersion of Earnings

Previous work shows that the self-employed have a wider dispersion of earnings than salaried workers, suggesting that self-employment is much riskier than salaried employment. However, the wider dispersion of earnings among the self-employed might not reflect greater earnings risk associated with entrepreneurship. In particular, past work mixes together two very heterogeneous groups of self-employed—incorporated and unincorporated business owners. The between group differences might account for the wider dispersion of earnings among the self-employed. Another, not mutually exclusive, explanation is that the wider dispersion reflects the heterogeneity among the incorporated and unincorporated self-employed above and beyond the gains and losses associated with self-employment. Thus, in this subsection, we examine the within group dispersion of earning gains.

Figures I and II report the quantile regression coefficient estimates from equation (3) for annual earnings for the incorporated and unincorporated respectively.[32] The gray bars provide the estimates without individual effects and are based on the specification used in column (5) of Table IX. The black bars provide the estimates with individual effects and are based on the specification in column (6) of Table IX. For the estimates without individual effects, we compare the difference between residual earnings of the incorporated self-employed at, for example, the 90th percentile of their earnings distribution with those of salaried workers at the 90th percentile of their distribution in Figure I. As demonstrated by the findings in Table IX, however, much of this gap, on average and at the median, reflects person specific factors. Thus, we also provide the quantile regression coefficient estimates when controlling for individual fixed effects in the black bars in

32. The figures are almost identical when we use equation (7) to control for selection within self-employment spells. Furthermore, when examining hourly earnings, the figures are similar except that the absolute values of the estimates are smaller at each quantile, reflecting the finding that people work more hours when self-employed.
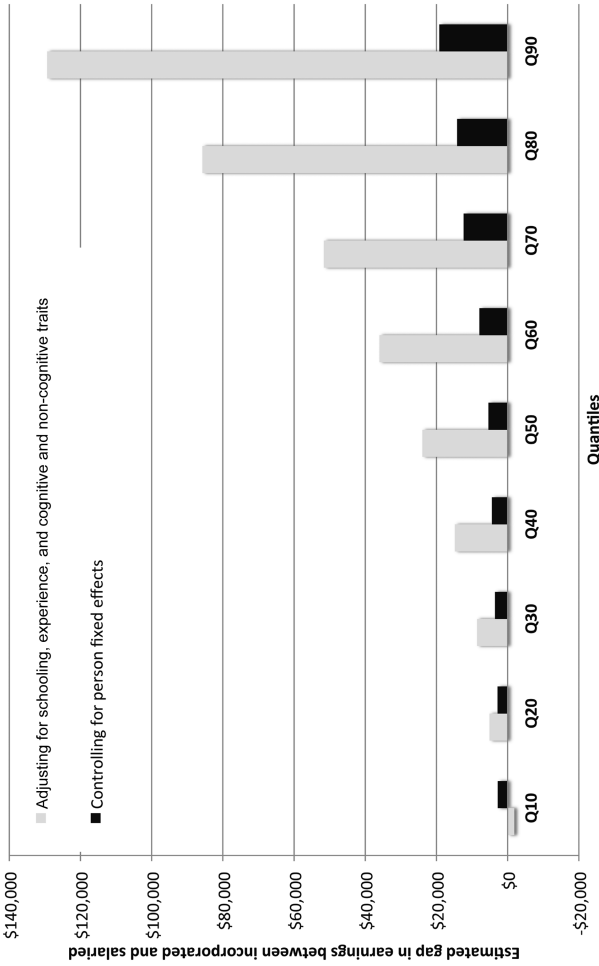
FIGURE I

Annual Earnings Gap Between Incorporated and Salaried

This figure depicts the quantile regression coefficient estimates of annual earnings on incorporated self-employment as specified in equation (3), where the estimates are provided for the 10th quantile (Q10), the 20th quantile (Q20), etc. The gray bars provide estimates while controlling for education (six categories), potential experience (quartic), AFQT, Rosenberg self-esteem, Rotter locus of control, and the Illicit Activity Index. The black bars provide estimates while also controlling for individual fixed effects. The sample includes white males, full-time workers, aged 25 years or older, as in Table IX.
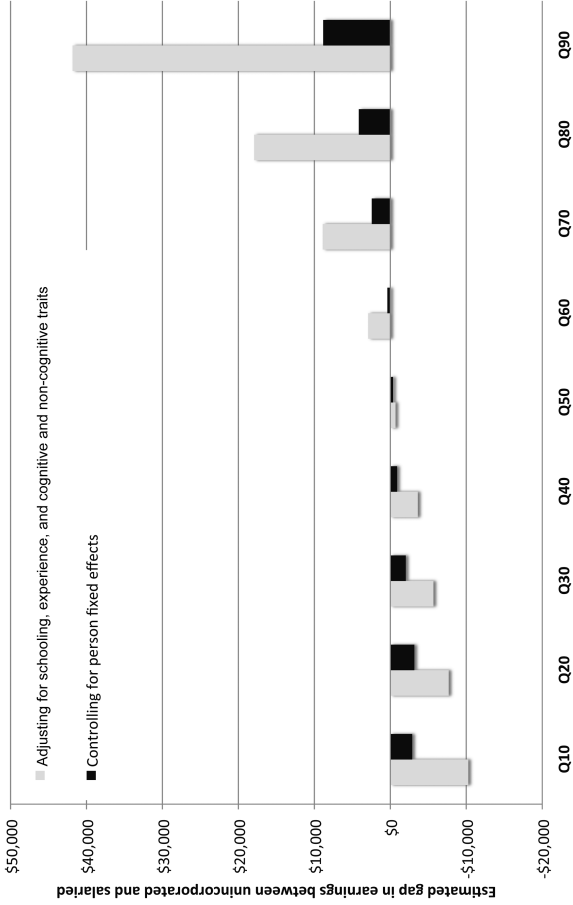
FIGURE II

Annual Earnings Gap between Unincorporated and Salaried

This figure depicts the quantile regression coefficient estimates of annual earnings on unincorporated self-employment as specified in equation (3), where the estimates are provided for the 10th quantile (Q10), the 20th quantile (Q20), etc. The gray bars provide estimates while controlling for education (six categories), potential experience (quartic), AFQT, Rosenberg self-esteem, Rotter locus of control, and the Illicit Activity Index. The black bars provide estimates while also controlling for individual fixed effects. The sample includes white males, full-time workers, aged 25 years or older, as in Table IX.

Figures I and II. When conducting quantile analyses with individual fixed effects, we shed light on the distribution of gains in earnings associated with self-employment for those who switched between salaried jobs and self-employment.

Figures I and II show that both the incorporated and unincorporated self-employed have wider dispersions of earnings than those of salaried workers with comparable traits. To see this, consider the gray bars. Figure I indicates that exceptionally successful incorporated business owners (90th percentile) tend to earn almost $130,000 more per annum than exceptionally successful salaried workers. Furthermore, notice that the estimated gap in residual annual earnings is positive from the 20th percentile onward. Most people who run incorporated businesses earn more, and for much of the distribution much more, than comparable salaried workers. But this is not true for the unincorporated; most of the unincorporated earn less, and some earn much less.

Figures I and II also show that individual effects account for much of the wider dispersion of the earnings for both the incorporated and unincorporated self-employed relative to salaried workers. That is, when we examine the gain in earnings associated with incorporated self-employment (Figure I) and unincorporated self-employment (Figure II) using within-person estimates (online, red bars), we find a much smaller dispersion in earnings than when we compare the earnings of salaried workers and those of the incorporated and unincorporated self-employed respectively (online, gray bars). For the incorporated self-employed, the estimated gain in earnings is positive at each decile, indicating that earnings tend to rise when individuals become incorporated self-employed across virtually the entire distribution. At the 90th percentile of the within-person gain in earnings associated with incorporated self-employment, a person tends to enjoy an almost $20,000 increase in earnings when becoming incorporated self-employed. For those who become unincorporated self-employed, the results are more nuanced. The within person gain in earnings associated with unincorporated self-employment is negative for more than half of the distribution, only becoming materially positive at the 70th percentile.[33]

33. We further assess the relationship between risk and employment type by examining the coefficient of variation of earnings. For the NLSY79 sample, we compute the coefficient of variation over employment spells. As shown in the Online Appendix Table XII, the coefficient of variation in earnings is greater when

VI. CONCLUSIONS

We disaggregate the self-employed into the incorporated and unincorporated to distinguish between "entrepreneurs" and other business owners. We show that incorporated business owners tend to engage in jobs that demand stronger nonroutine cognitive skills than either unincorporated business owners or salaried workers. In contrast, unincorporated business owners tend to perform tasks that demand comparatively strong manual skills. To the extent that one associates entrepreneurship with analytical reasoning, creativity, and complex interpersonal communications rather than with eye, hand, and foot coordination, the data suggest that on average the incorporated self-employed engage in entrepreneurial activities while the unincorporated do not.

We discover that entrepreneurs—as proxied by the incorporated self-employed—earn more than and have a very distinct mixture of cognitive and noncognitive traits from salaried workers and other business owners. The incorporated tend to be male, white, better educated, and are more likely to come from high-earning, two-parent families. Furthermore, as teenagers, the incorporated tend to have higher learning aptitude and self-esteem scores. But, apparently it takes more to be a successful entrepreneur than having these strong labor market skills: the incorporated self-employed also tend to engage in more illicit activities as youths than other people who succeed as salaried workers. It is a particular mixture of traits that seems to matter for both becoming an entrepreneur and succeeding as an entrepreneur. It is the high-ability person who tends to break the rules as a youth who is especially likely to become a successful entrepreneur.

UNIVERSITY OF CALIFORNIA, BERKELEY, AND NBER
LONDON SCHOOL OF ECONOMICS, CENTRE FOR ECONOMIC PERFORMANCE, AND CEPR

SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online.

---

a person is an incorporated business owner than when the person is a salaried worker. However, the coefficient of variation in earnings among the incorporated self-employed is much lower than the S&P 500 or long-term government bonds, and the average boost in earnings associated with becoming an incorporated business owner is larger than the boost in earnings associated with shifting from short-term T-bills to the S&P 500 or longer-term government bonds.

REFERENCES

Autor, David H., Frank Levy, and Richard J. Murnane, "The Skill Content of Recent Technological Change: An Empirical Investigation," *Quarterly Journal of Economics*, 118 (2003), 1279–1333.

Baumol, William J., "Entrepreneurship: Productive, Unproductive, and Destructive," *Journal of Political Economy*, 98 (1990), 893–921.

Bengtsson, Ola, Tino Sanandaji, and Magnus Johannesson, "Do Women Have a Less Entrepreneurial Personality?," Research Institute of Industrial Economics Working Paper No. 944, 2012.

Bernardo, Antonio E., and Ivo Welch, "On the Evolution of Overconfidence and Entrepreneurs," *Journal of Economics and Management Strategy*, 10 (2001), 301–330.

Bertrand, Marianne, and Antoinette Schoar, "Managing with Style: The Effect of Managers on Firm Policies," *Quarterly Journal of Economics*, 118 (2003), 1169–1208.

Blanchflower, David G., and Andrew J. Oswald, "What Makes an Entrepreneur?," *Journal of Labor Economics*, 16 (1998), 26–60.

Bloom, Nicholas, and John Van Reenen, "Measuring and Explaining Management Practices across Firms and Countries," *Quarterly Journal of Economics*, 122 (2007), 1351–1408.

Borjas, George J., and Stephen G. Bronars, "Consumer Discrimination and Self-Employment," *Journal of Political Economy*, 97 (1989), 581–605.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne, "The Determinants of Earnings: A Behavioral Approach," *Journal of Economic Literature*, 39 (2001), 1137–1176.

Budig, Michelle J., "Intersections on the Road to Self-Employment: Gender, Family and Occupational Class," *Social Forces*, 84 (2006), 2223–2239.

Carr, Deborah, "Two Paths to Self-Employment?: Women's and Men's Self-Employment in the United States, 1980," *Work and Occupations*, 23 (1996), 26–53.

Dawson, Christopher, David de Meza, Andrew Henley, and G. Reza Arabsheibani, "Entrepreneurship: Cause or Consequence of Financial Optimism?," *Journal of Economics and Management Strategy*, 23 (2014), 717–742.

De Meza, David, and C Southey, "The Borrower's Curse: Optimism, Finance, and Entrepreneurship," *Economic Journal*, 106 (1996), 375–386.

Evans, David S., and Boyan Jovanovic, "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy*, 97 (1989), 808–827.

Evans, David S., and Linda S. Leighton, "Some Empirical Aspects of Entrepreneurship," *American Economic Review*, 79 (1989), 519–535.

Fairlie, Robert W., "Drug Dealing and Legitimate Self-Employment," *Journal of Labor Economics*, 20 (2002), 538–567.

———, "Self-Employment, Entrepreneurship, and the NLSY79," *Monthly Labor Review*, 128 (2005), 40–47.

Fairlie, Robert W., and Bruce D. Meyer, "Ethnic and Racial Self-Employment Differences and Possible Explanations," *Journal of Human Resources*, 31 (1996), 757–793.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, "Human Capital and Regional Development," *Quarterly Journal of Economics*, 128 (2013), 105–164.

Glaeser, Edward L., "Entrepreneurship and the City," NBER Working Paper No. w13551, 2007.

Hamilton, Barton H., "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment," *Journal of Political Economy*, 108 (2000), 604–631.

Hartog, Joop, Mirjam van Praag, and Justin van derSluis, "If You Are So Smart, Why Aren't You an Entrepreneur? Returns to Cognitive and Social Ability:

Entrepreneurs versus Employees," *Journal of Economics and Management Strategy*, 19 (2010), 947–989

Heckman, James J., "Policies to Foster Human Capital," *Research in Economics*, 54 (2000), 3–56.

Heckman, James J., and Yona Rubinstein, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *American Economic Review*, 91 (2001), 145–149.

Heckman, James J., N. Stixrud, and S. Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24 (2006), 411–482.

Hipple, Steven F., "Self-Employment in the United States," *Monthly Labor Review*, 113 (2010), 17–32.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen, "Sticking it Out: Entrepreneurial Survival and Liquidity Constraints," *Journal of Political Economy*, 102 (1994), 53–75.

Horvath, Paula, and Marvin Zuckerman, "Sensation Seeking, Risk Appraisal, and Risky Behavior," *Personality and Individual Differences*, 14 (1993), 41–52.

Hurst, Erik, Geng Li, and Benjamin W. Pugsley, "Are Household Surveys Like Tax Forms: Evidence from Income Underreporting of the Self-Employed," *Review of Economics and Statistics*, 96 (2014), 19–33.

Hurst, Erik, and Benjamin W. Pugsley, "What Do Small Businesses Do?," *Brookings Papers on Economic Activity* (2011), 73–118.

Knight, Frank H. *"Risk, Uncertainty and Profit* (New York: Harper and Row, 1921).

Kushner, David, "Machine Politics: The Man Who Started the Hacker Wars," *New Yorker*, May 7, 2012; retrieved from http://www.newyorker.com/magazine/2012/05/07/machine-politics.

La Porta, Rafael, and Andrei Shleifer, "The Unofficial Economy and Economic Development," *Brookings Papers on Economic Activity*, (2008), 275–352.

———, "Informality and Development," *Journal of Economic Perspectives*, 28 (2014), 109–126.

Lazear, Edward P., "Balanced Skills and Entrepreneurship," *American Economic Review*, 94 (2004), 208–211.

———, "Entrepreneurship," *Journal of Labor Economics*, 23 (2005), 649–680.

Lucas, Robert E., "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9 (1978), 508–523.

Madrian, Bridgette C., and Lars John Lefgren, "An Approach to Longitudinally Matching Current Population Survey (CPS) Respondents," *Journal of Economic and Social Measurement*, 26 (2000), 31–62.

Malmendier, Ulrike, and Geoffrey Tate, "Superstar CEOs," *Quarterly Journal of Economics*, 124 (2009), 1593–1638.

Manso, Gustavo, "Experimentation and the Returns to Entrepreneurship," *Review of Financial Studies*, forthcoming (2016).

Mohapatra, Sandeep, Scott Rozelle, and Rachael Goodhue, "The Rise of Self-Employment in Rural China: Development or Distress?," *World Development*, 35 (2007), 163–181.

Moskowitz, Tobias, and Annette Vissing-Jørgensen, "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?," *American Economic Review*, 92 (2002), 745–778.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, "The Allocation of Talent: Implications for Growth," *Quarterly Journal of Economics*, 106 (1991), 503–530.

Nicolaou, Nicos, Scott Shane, Lynn Cherkas, and Tim D. Spector, "The Influence of Sensation Seeking in the Heritability of Entrepreneurship," *Strategic Entrepreneurship Journal*, 2 (2008), 7–21.

Özcan, Berkay, "Only the Lonely? The Influence of the Spouse on the Transition to Self-Employment," *Small Business Economics*, 37 (2011), 465–492.

Queiro, Francisco, "The Effect of Managerial Education on Firm Growth," Mimeo, Harvard University, 2015.

Schumpeter, Joseph A., *Theorie Der Wirtschaftlichen Entwicklung* (Leipzig: Duncker & Humblot, 1911).

Smith, Adam, *An Inquiry into the Nature and Causes of the Wealth of Nations* (London: Methuen, 1776).

Zuckerman, Marvin, *Behavioral Expressions and Biosocial Bases of Sensation Seeking* (New York: Cambridge University Press, 1994).