

The Right Stuff? Personality and Entrepreneurship*

Barton H. Hamilton[†]

Olin Business School, Washington University in St. Louis

Nicholas W. Papageorge[‡]

Department of Economics, Johns Hopkins University, IZA and NBER

Nidhi Pande[§]

Department of Economics, University of Delhi

September 21, 2018

ABSTRACT: We construct a structural model of entry into self-employment to evaluate the impact of policies supporting entrepreneurship. Previous work has recognized that workers may opt for self-employment due to the non-pecuniary benefits of running a business and not necessarily because they are good at it. Other literature has examined how socio-emotional skills, such as personality traits, affect selection into self-employment. We link these two lines of inquiry. The model we estimate captures three factors that affect selection into self-employment: credit constraints, relative earnings and preferences. We incorporate personality traits by allowing them to affect sector-specific earnings as well as preferences. The estimated model reveals that the personality traits that make entrepreneurship profitable are not always the same traits driving people to open a business. This has important consequences for entrepreneurship policies. For example, subsidies for small businesses do not attract talented-but-reluctant entrepreneurs, but instead attract individuals with personality traits associated with strong preferences for running a business and low-quality business ideas.

KEYWORDS: Entrepreneurship, Personality, Socio-emotional skills, Latent factors.

JEL CLASSIFICATION: J23, J24, J31, J32

*We gratefully acknowledge helpful comments from: Thomas Åstebro, Jorge Balat, Robert Fairlie, George Levi-Gayle, Bruce Hall, Bruce Hamilton, Mitchell Hoffman, Andrew Knight, Robert Pollak, Victor Ronda, Yuya Sasaki, Kathryn Shaw, Richard Spady and Matthew Wiswall along with seminar participants at Chinese University in Hong Kong, The Hong Kong University of Science and Technology, University College London, Cambridge University, University of Essex, Georgia Tech, the 2013 Conference on New Directions in Applied Microeconomics at Cal Tech, the 2014 SOLE Meetings, the 2014 European and North American Meetings of the Econometric Society, the Venice Summer Institute on the Economics of Entrepreneurship and the 6th IZA/Kauffman Foundation Workshop on Entrepreneurship Research. We also thank the editor and two anonymous referees for helpful guidance on this paper. The usual caveats apply.

[†]hamiltonb@wustl.edu.

[‡]papageorge@jhu.edu.

[§]nidhi13pande@gmail.com.

1 Introduction

Entrepreneurship has occupied economic thought for nearly a century. This sustained interest reflects a widely-held view that individuals pursuing their own business ventures drive innovation and economic growth (Schumpeter, 1949). Entrepreneurship, however, remains poorly understood. Most small businesses fail, but it is unclear why some individuals are successful entrepreneurs while others are not. Even more puzzling is evidence showing that most individuals who remain self-employed would earn more in traditional paid employment (Hamilton, 2000). Recent research in economics has led to the acknowledgement of the role of socio-emotional, non-cognitive or *soft* skills — including personality traits — in driving economic behavior like labor supply.¹ This shift raises the question: could personality differences explain which individuals become entrepreneurs and, among those who enter, which ones succeed?²

We examine how socio-emotional skills affect both entry into self-employment and entrepreneurial returns. To measure socio-emotional skills, we use the *Big 5* personality traits, which will be discussed in detail in Section 2. We estimate a model in which agents who face credit constraints maximize utility by choosing between self- and paid employment. Previous literature has recognized the possibility that workers opt for self-employment because they enjoy it and not because they are good at it.³ Other research has demonstrated how entrepreneurs differ from paid employees on a variety of important dimensions, including socio-emotional skills (Levine and Rubinstein, 2017). The model we specify links these two lines of inquiry by distinguishing between the roles of sector preferences and sector performance in determining entry, where personality is allowed to affect both. We also exploit multiple measures of personality taken over the lifecycle to identify the distributions of latent, stable and possibly correlated underlying traits, thus circumventing possible mis-measurement issues associated with standard personality assessments. Using our setup, we obtain sector-specific market prices of latent personality traits along with estimates of how personality links to preferences over sectors.

We highlight two key features of our model, both of which are essential for assessing

¹Economists have yet to settle on the nomenclature. In this paper, we focus on “personality traits” which we sometimes refer to collectively as “personality”. In our discussion, we view personality traits as a subset of “non-cognitive” or “socio-emotional” skills.

²In this study, we define an entrepreneur as an individual who reports self-employment.

³For example, Hamilton (2000) shows evidence of non-pecuniary benefits to self-employment, while Hurst and Pugsley (2011) use data from a survey to show that most new small business owners do not plan to grow very much, but do report strong non-pecuniary benefits of being their own boss. Our work complements these studies. One difference from the latter piece is that we rely on revealed preferences versus stated intentions. We also construct a structural model of entry that can be used to evaluate policy given how preferences and expected earnings affect the decision to become self-employed.

counterfactual policies, such as subsidies. First, the model captures various mechanisms affecting entry into self-employment. Capturing selection is crucial since policies such as subsidies shift the composition of individuals who sort into self-employment and thus the quality of businesses that are started. In the model, selection arises due to relative earnings in paid employment, credit constraints and preferences. For example, a “lifestyle entrepreneur” may choose to open a business based on a low-quality idea since his personality means he enjoys the autonomy of being his own boss. Alternatively, what we term a “reluctant entrepreneur” may have a personality type that is productive in self-employment, but also predicts an aversion to being an entrepreneur. These types of misalignments can influence the impact of policies designed to promote entrepreneurship. Subsidies might be useful if they induce talented but reluctant entrepreneurs into self-employment, but could be wasteful if they simply attract lifestyle entrepreneurs to opening unprofitable businesses.

A second important feature of our model, which follows Evans and Jovanovic (1989) but departs from many prior studies, is that agents are assumed to observe the quality of their business idea prior to choosing whether to open a business.⁴ This is in contrast to models where agents are assumed to lack knowledge about the business they would open and instead choose a sector based on average earnings differences across sectors (Willis and Rosen, 1979; Rees and Shah, 1986).⁵ This approach is perhaps defensible if mean earnings approximate median earnings. However, given the highly right-skewed self-employment earnings distribution, averages in the context we study vastly exceed what nearly all potential entrepreneurs can expect to earn. Using mean earnings as expected earnings is thus potentially misleading. Doing so can generate the erroneous conclusion that there is a mass of reluctant potential entrepreneurs forgoing high expected earnings in self-employment and who thus have a distaste for opening a business which could be overcome through policies, such as a subsidy. Our approach is to model the worker’s information set to include the expected value of his own potential business idea. Our modeling assumption is supported by recent research suggesting that individuals opening businesses are aware of the quality of their venture prior to entry. For example, Levine and Rubinstein (2017) and Hincapié (2018) show that individuals who start more successful businesses make the costly effort of incorporating their businesses prior to the earnings realization.

We estimate the model using data from the 1995 and 2004 waves of the National Survey of Midlife Development in the United States (MIDUS). Estimates reveal that the personality traits that make entrepreneurship most profitable are not the same personality traits that

⁴As described below, we model entrepreneurial income as a function of the quality of the business idea, capital, and a shock that is not known by the agent *ex ante*.

⁵Some recent papers relax the limited information assumption by allowing for individual types, which are known by the individual prior to entry. See, e.g., Humphries (2017).

drive people to open their own business. For example, similar to earlier work (see, for example, Caliendo, Fossen, and Kritikos (2014)), we find evidence that two of the Big 5 personality traits, extraversion and “openness to new experiences,” predict higher rates of self-employment. However, since we explicitly distinguish between preferences versus performance to explain the entry decision, we can go beyond earlier work to isolate different reasons why. We show that extraverted individuals are attracted to entrepreneurship because they earn more in self-employment than in paid employment. In contrast, open individuals perform poorly in self-employment, but exhibit a strong preference for starting a business which offsets their low expected earnings enough to induce entry. Identifying this type of misalignment links our work to the more general idea that socio-emotional skills can have different impacts in different sectors. This point is often overlooked in the literature, though a notable exception is Lundberg (2013), which shows that the role of personality in predicting educational attainment varies by socio-demographic group.⁶ Capturing this type of misalignment also allows us to understand their consequences for counterfactual policies.

Using our estimated model, we simulate decisions and returns under a counterfactual policy removing credit constraints. Doing so, we show that credit constraints do not prevent good ideas from entering the market (and may even screen out a small number of low-quality ideas). However, removing constraints can prevent some businesses from operating at sub-optimally small scales. We also show that counterfactual subsidies are largely ineffective. One reason is that they subsidize businesses that would have been started absent support. Such payments also attract individuals into entrepreneurship who possess traits, such as openness to new experiences, which are associated with strong preferences for, but weak performance in, self-employment. The result of these policies is an increase in entry but a decline in the average pecuniary value of realized business ideas. These findings suggest that policies that encourage entrepreneurship are potentially wasteful.⁷

This study contributes to three separate literatures. The first studies the decision to open a business. In a seminal paper, Evans and Jovanovic (1989) show that credit constraints

⁶Other papers include that by Lundberg (2012), who shows that the pecuniary returns to personality factors vary both by tenure and by educational group, suggesting that different personality traits may enhance productivity in some occupations, but not others. Levine and Rubinstein (2017) show that deviant behavior can lead to successful entrepreneurship and Papageorge, Ronda, and Zheng (2016) show that some forms of childhood misbehavior capture socio-emotional skills that predict higher earnings despite also being associated with lower educational attainment. Prada and Urzúa (2016) show that mechanical skill can reduce four-year college attendance — not necessarily due to low academic ability, but instead due to high returns in the labor market conditional on not attaining a four-year degree. See also Almlund et al. (2011), who stress the importance of accounting for varying returns to socio-emotional skills and Cattan (2011), who develops this point for traits related to an individual’s self-confidence and attitudes towards women.

⁷Relatedly, Hurst and Pugsley (2014) provide a theoretical model of entrepreneurship that includes non-pecuniary benefits. Their model predicts that some policies promoting self-employment can be distortionary.

are binding so that individuals with especially profitable ideas, but few assets, may be unable to pursue their business venture. Relatedly, Paulson, Townsend, and Karaivanov (2006) show that credit constraints alone cannot explain why good business ideas are not pursued and that moral hazard also plays a role. Both papers suggest that some paid employees would be successful entrepreneurs were it not for market imperfections. On the other hand, Hamilton (2000) shows that many entrepreneurs who are “successful” in that their businesses have not failed would have earned more had they remained in traditional paid employment. This finding may reflect important non-pecuniary benefits to self-employment, such as autonomy.⁸ Taken together, this research leads to the following somewhat startling conclusion: entrepreneurship does not necessarily attract the subset of individuals for whom it would generate the highest pecuniary returns.

A second related literature, much of it from personnel psychology, studies how measurements of personality traits relate to job performance and job satisfaction. Barrick and Mount (1991) show that individuals who are open to new experiences are especially good trainees, perhaps since they are eager to try new things. However, they are not necessarily better employees. More closely related to self-employment, Barrick and Mount (1993) show that two other traits, conscientiousness and extraversion, are associated with better job performance, especially for managers who exercise more autonomy at work. Since autonomy is a hallmark of self-employment, this finding suggests that the relationship between personality and success differs in paid versus self-employment.⁹ Further work from psychology has directly examined how self-employment and personality are connected, suggesting, for example, that entrepreneurs score highly on the trait openness to new experiences, which is generally consistent with our findings.¹⁰

A third, burgeoning literature to which we contribute incorporates socio-emotional skills and personality traits into economic models of rational decision-making. Much of this work

⁸In another key contribution, Lazear (2004) shows that a successful entrepreneur must be a “jack-of-all-trades” with a wide variety of skills. Our focus is different in that we examine how a fixed set of skills affect entrepreneurial entry and returns, whereas Lazear (2004) considers skills that are acquired or learned through optimal investments. Fairlie and Holleran (2012) and Fairlie, Karlan, and Zinman (2015) connect these two ideas, showing that personality can affect short-run responsiveness to a training program for entrepreneurs (though they find no evidence of long-run effects of the program). Also related, Åstebro and Thompson (2011) argue that entrepreneurs acquire a range of skills in part due to preferences for variety.

⁹From economics, Cubel et al. (2016) assess the relationship between personality traits and productivity. They circumvent selection issues by measuring productivity in a laboratory setting. They demonstrate that more conscientious people perform better and more neurotic people perform worse. Although we use observational data, we believe our study complements their research since we also aim to address how personality can affect both selection into sectors and sector-specific performance.

¹⁰These analyses include: Hisrich, Langan-Fox, and Grant (2007), Zhao and Seibert (2006), Brandstätter (2011), Zhao, Seibert, and Lumpkin (2010) and Rauch and Frese (2007).

can be traced to Heckman and Rubinstein (2001).¹¹ They show that socio-emotional skills can account for much of the observed variance in sociodemographic outcomes. Building on this work, economists have studied how personality traits and socio-emotional skills relate to a host of outcomes, including marriage (Lundberg, 2012, 2011), education (Barón and Cobb-Clark, 2010; Savelyev, 2010; Gensowski, Heckman, and Savelyev, 2011; Heckman and LaFontaine, 2010; Heckman, Pinto, and Savelyev, 2013) and health (Heckman, 2012). More closely related to our study are papers relating personality to labor market behavior (Heckman, Stixrud, and Urzua, 2006; Urzua, 2008; Wichert and Pohlmeier, 2009; Heineck, 2010; Störmer and Fahr, 2013). This research has led to some particularly striking results, showing, for example, that socio-emotional skill differences can help explain education and earnings differences between men and women or between black and white individuals.

Comparatively little research in economics has directly connected self-employment and socio-emotional skills. Notable exceptions include the aforementioned study by Caliendo, Fossen, and Kritikos (2014); Levine and Rubinstein (2017), who show evidence that entrepreneurs differ from paid employees on a number of socio-emotional dimensions; Hartog, Van Praag, and Van Der Sluis (2010), who examine “social ability” and entrepreneurial firms; Asoni (2010), who studies self-employment spells and self-confidence; and Humphries (2017), who studies a uni-dimensional measure of non-cognitive skill coming from a test administered to Swedish army recruits. A key departure for this study is to explicitly link multiple dimensions of socio-emotional skill (five personality traits) to various features of selection and performance in self-employment, thus increasing the set of conclusions we can draw about potential policies affecting the entry decision.

The remainder of the paper proceeds as follows. Section 2 introduces the “Big 5” personality traits. Section 3 discusses the data we use. Section 4 specifies the model and Section 5 discusses estimation. Section 6 examines parameter estimates. Section 7 presents results from counterfactual experiments. Section 8 concludes.

2 The “Big Five” Personality Traits

A large literature in psychology has settled upon five traits (the Big 5), which summarize an individual’s personality. These five are chosen using statistical models (often known as factor models) intended to focus attention on traits that are neither overlapping nor redundant. As with any rubric, there is some debate surrounding the Big 5, but they are attractive for a

¹¹Excellent summaries of the state of this line of research are found in Borghans et al. (2008) and Almlund et al. (2011). The techniques used in this literature draw upon Goldberger (1972) and Jöreskog and Goldberger (1975).

few reasons.¹² While research on the technology of skill formation points to the mutability of character for children and adolescents (Cunha, Heckman, and Schennach, 2010; Heckman and Kautz, 2013), personality traits appear to be relatively stable over the adult lifecycle (Caspi, 2000; Cobb-Clark and Schurer, 2012). One explanation for stability comes from evidence using data on twins suggesting a genetic basis for personality traits (Zhang et al., 2009; Shane et al., 2010; Shane and Nicolaou, 2013). The stability of personality traits among adults should dispel concern about simultaneity if the Big 5 are used as right-hand-side variables in regressions explaining economic behavior. As described below, we investigate this issue by exploiting multiple assessments of a given individual’s personality in our data to show that self-employment and earnings do not predict personality traits measured later in life.¹³ Originally proposed in Goldberg (1971), the Big 5 are: *agreeableness*, *extraversion*, *neuroticism*, *conscientiousness* and *openness to new experiences*. The characteristics used to measure them are listed in Table 1.

Despite the growing and fruitful integration of personality measures into economic models, important conceptual problems remain (Almlund et al., 2011). Most problematic is how (or even whether) personality fits within the utility paradigm in economics. Personality traits may reflect or be correlated with preferences. Alternatively, as Almlund et al. (2011) propose, personality and preferences may both reflect some deeper, as yet unknown characteristic, which drives human behavior. Some recent work addresses this issue, proposing models that explicitly link preferences with socio-emotional skills (Bowles, Gintis, and Osborne, 2001; Rustichini et al., 2016). Bowles, Gintis, and Osborne (2001), for example, model personality as enhancing preferences. Other researchers have used laboratory experiments to ascertain how socio-emotional abilities relate to measures more familiar to economists, including preferences over risk, time and ambiguity (Dohmen et al., 2008, 2010; Fréchet, Schotter, and Trevino, 2017; Vandenberghe, St-Onge et al., 2008).¹⁴

One way forward is to think of personality as affecting the utility cost of time in different activities. If we accept that hours spent in each employment sector imply a distinct utility cost, our model effectively suggests that sector-specific utility costs can differ by personality. Agents with different personalities will then differ in their sector choices once we have controlled for differences in pecuniary returns in each sector. Personality traits may also affect the amount of effort or time used to produce a given amount of output in each sector so that the opportunity costs of production differ by personality traits in self- versus paid

¹²Some rubrics suggest a sixth trait, which seems to capture agency or control. We focus on the Big 5 as it is the most common rubric.

¹³These findings are discussed in Section 5, when we discuss estimation.

¹⁴Further work on issues integrating personality into economics is found in Heckman and Kautz (2012), Roberts et al. (2011) and Borghans et al. (2011).

employment. This thinking would align our model with the framework proposed in Becker (1965), who emphasizes that preferences over consumption reflect how different goods take different amounts of time to consume.

3 Data

In this section, we conduct a preliminary analysis of the data set used in the paper, the National Survey of Midlife Development in the United States (MIDUS), and highlight two empirical patterns. First, we show that individuals with more assets are more likely to be self-employed. However, conditional on self-employment, there is little evidence that individuals with more assets have more profitable business ideas. Second, we illustrate the idea that some personality traits can have opposing effects on preferences versus relative performance in each sector, showing that “openness to new experiences” simultaneously predicts entry into self-employment and relatively low earnings in self-employment.

3.1 The MIDUS Data Set

The MIDUS survey studies midlife from an unusually rich variety of perspectives. Information is collected on the labor market choices and outcomes, physical health, and psychological well-being of a representative sample of working age men and women in the United States. Crucially for the present study, the MIDUS data set includes information on whether individuals are self-employed, their assets and standard measures of the Big 5 personality traits.¹⁵

MIDUS data collection occurred in two waves, the first (MIDUS I) in 1995 and the second (MIDUS II) in 2004. The sample surveyed in 1995 included over 7,114 men and women between ages 25 and 74 from the United States. The second wave surveyed a nationally representative subsample of 4,009 individuals with the goal of understanding the physical, health and psychological effects of aging. In our study, we use both waves of data, including each individual’s answers on two personality assessments. Using both assessments helps us to circumvent possibly mis-measured personality traits (including the effect of aging on responses to personality assessments for a given latent factor). In particular, we use multiple

¹⁵To our knowledge, only two previous papers in economics make use of the MIDUS data set. They are Lundborg (2013) and Cutler and Lleras-Muney (2010). The MIDUS survey was administered by the John D. and Catherine T. MacArthur Foundation Research Network on Successful Midlife Development. The survey is designed to be nationally representative, but over-weights older men to better assess midlife (MIDMAC, 1999).

measures to identify the distribution of permanent latent factors that are measured by the personality assessments.

The MIDUS data collects information on household assets in 1995.¹⁶ Notice, this measure includes salable assets, which thus excludes pensions, retirement wealth or expected social security income, which would be difficult or illegal to sell to invest in a business. In our reduced form analysis, we use 1995 assets as a proxy for 2003 assets, which are relevant for business investments. We address this type of measurement error more formally in our structural estimation detailed in Sections 4 and 5

In constructing our analytic sample, we restrict attention to male workers that are under age 65 in 2004. Starting with the original sample of 7,114, by focussing on men, we drop 3,719 observations. By focussing on working-age men we lose another 449 observations. We also drop individuals who are not working, losing another 231 observations. By dropping observations with missing data on assets, we lose 550. Finally, we drop an additional 279 observations missing information on other key variables. Of the remaining 1,886 observations, 990 are not observed in 2004, which leaves us with an analytic sample of 896 working men in 2004 who have full information on key variables, including income, sector choice, personality traits and assets in 1995.¹⁷

Summary statistics are found in Table 2 for the analytic sample of 896 men observed working in 2004 and then separately for individuals in self- versus paid employment. We also include differences in means between these two groups and p -values from t-tests of

¹⁶Individuals are asked “Suppose you (and your spouse or partner) cashed in all your checking and savings accounts, stocks and bonds, real estate, sold your home, your vehicles, and all your valuable possessions. Then suppose you put that money toward paying off your mortgage and all your other loans, debts, and credit cards. Would you have any money left over after paying your debts or would you still owe money?” Individuals then report the amount of assets in bins. From \$1-\$19,999, increments are \$1,000; from \$20,000-\$99,999, increments are \$5,000; from \$100,000 to \$499,999, increments are \$50,000. Remaining increments are \$500,000-\$999,999 and \$1,000,000 or more. Individuals are assigned the midpoint of the bin they report and those who have negative net assets are assigned 0 salable assets.

¹⁷A concern with our use of data from MIDUS II is that the sample is selected towards individuals who were located for a follow up and agreed to participate again. Of the 990 individuals who are observed in 1995, but not in 2004, 138 are either not working (or are not working age) in 1995. Thus, we observe 852 working men in 1995 who are not observed in 2004. Compared to these 852 men, the men in the analytic sample of 896 workers observed in 2004 report higher assets, earned more in 1995 and are more highly educated in 1995. Personality traits are generally not significantly different across these samples, but individuals who are not observed in 2004 are more open to new experiences compared to those who did respond. There are no differences across these two groups in the 1995 likelihood of self-employment. A concern arises if missing individuals are selected on unobservables, which limits the internal validity of estimates, especially those related to asset growth between 1995 and 2003. How we correct for this in our model and the various tests for robustness we perform are discussed in Sections 4 and 5. Missing data is also a concern for external validity if individuals in our sample exhibit relationships between personality and self-employment that are not representative of the population. Reassuringly, extraversion and openness predict self-employment in our sample, which is consistent with other papers relating personality to entrepreneurship.

whether group differences are significant.¹⁸ According to the table, entrepreneurs earn more on average than paid employees. One explanation is that entrepreneurship is more lucrative than paid employment. However, as Hamilton (2000) points out, these types of averages ignore selection into sectors. A high-earning entrepreneur may have earned the same or more had they chosen paid employment. Moreover, averages obscure the skewed distribution of earnings. Looking at medians, we find that the typical individual would not expect to earn more by opening his own business. The model specified in Section 4 incorporates both relative sector earnings and the skewed distribution of self-employment income.

Table 2 reports average sociodemographic and personality measures. We find that education, marriage, number of children and spouse’s education do not differ systematically by sectors.¹⁹ However, conditional on marriage, for individuals who choose self-employment, spousal education is slightly higher as is the likelihood of a spouse being employed (though the latter is only significant at the 11% level).²⁰ Cognitive skill, as measured by *fluid cognitive ability*, is likewise the same across sectors. However, we do find average differences by sector in the Big 5 personality traits. For example, Table 2 shows that entrepreneurs tend to be more agreeable, extraverted and open to new experiences than paid employees. The latter two traits are typically associated with entrepreneurship.

3.2 Assets and Self-Employment

Table 2 also shows that individuals in self-employment have, on average, about double the assets of paid employees. This suggests that individuals may select into self-employment based on their ability to fund their own business. To investigate this relationship further, Figure 1(a) plots self-employment entry against assets and includes a fitted non-parametric polynomial. The figure shows that much of the increase in self-employment by assets occurs at moderate levels of wealth (below \$200,000).

Credit constraints would suggest that, conditional on a business idea, assets would drive entry. An alternative explanation is that high-asset individuals are more productive in self-employment. To assess this possibility, Figure 1(a) plots self-employment earnings by

¹⁸Table S1, found in the appendix, provides summary statistics for a larger sample of all working males who participated in the second wave of MIDUS data collection but who may be missing information for some variables. We show that key patterns in the data are robust to the inclusion of these individuals.

¹⁹In comparison to 2004 Current Population Survey averages, MIDUS II participants report higher education.

²⁰There are many reasons why this might be the case, including the possibility of risk-sharing or access to benefits like subsidized health insurance. Entrepreneurs can effectively use their spouse’s more steady employment or benefits as a safety net given the high probability of failure and the lack of benefits typical in self-employment.

assets in 1995. Two patterns emerge. First, there is some evidence from the raw data that men with more assets earn more in self-employment, especially at very high levels (i.e., above \$500,000), which could mean that men with more assets expect to earn more in self-employment. Alternatively, it suggests that assets do not predict entry, but instead lead to under-investment of very high-quality business ideas, forcing potentially profitable enterprises to operate at sub-optimal scales. Second, and relatedly, credit constraints do not appear to bar low-asset individuals from opening lucrative businesses. Figure 1(b) shows a cluster of individuals with near-zero assets who open businesses that generate high earnings (on the order of \$50,000-\$100,000), suggesting that credit constraints may not restrict entry into self-employment. In light of these empirical patterns, our structural model will exploit data on assets to identify possible credit constraints that potential entrepreneurs face, which allows for the possibility of starting a business despite sub-optimal investments in the venture. Moreover, the production function we use to model how business ideas generate income will be specified so that low-asset individuals with profitable ideas are not precluded from entry through, for example, some minimum level of assets needed to go into business. This way, low asset individuals can potentially profit from very good ideas. Finally, the model accommodates potential correlation between earnings shocks, business ideas and assets. This accounts for the possibility that individuals with higher expected earnings in either self- or paid employment may have accumulated more assets, which they can invest in their business venture.

3.3 Openness, Earnings and Self-Employment

Prior studies consistently find a strong positive relationship between the personality trait “openness to new experiences” and the probability of self-employment. In Figure 1(c), we plot a binary variable for self-employment in 2004 against the 2004 measure of openness for our analytic sample. We add a smoothed polynomial fitted line with 95% confidence intervals. The figure shows that the probability of self-employment increases with openness. However, in Figure 1(d), we plot expected log earnings differences between self- and paid employment against 2004 openness. To do this, we first regress log earnings onto personality traits, cognition and a series of sociodemographic observables (age, education and marriage) separately by employment sector. Next, we use estimated coefficients to predict log earnings for each individual and sector, which we use to compute the expected sectoral difference (self minus paid). The result is a log earnings differential for each individual. We plot each individual’s self-versus-paid earnings differential against their 2004 openness score. We also plot a smoothed polynomial-fitted line along with 95% confidence intervals. The scatter plot

and fitted line show that the expected earnings premium in self-employment declines with openness. Moreover, the decline is both significant and monotonic. Together, Figures 1(c) and 1(d) provide preliminary empirical evidence that openness has mixed effects, predicting entry into a sector in which it generates relatively low returns.

The dueling effects of openness remain when we control for other variables that we expect to affect earnings and sector choices, including other personality traits and socio-demographic variables. Results from these regressions are presented in Table 3. Columns 1 and 2 report estimates from OLS regressions of log earnings in self- and paid employment, respectively. Sector-specific prices vary for a number of factors, including openness, where the coefficient in self-employment is -0.27 and in paid employment is 0.06.²¹ Column 3 presents probit estimates where the outcome variable is an indicator for self-employment. The estimates are similar to the estimates found in previous work linking entrepreneurship and personality (Caliendo, Fossen, and Kritikos, 2014). In particular, the coefficient on openness is positive and significant.²²

The finding in Table 3 that openness is associated with both a higher propensity for self-employment and lower self-employment earnings highlights the limitations of a reduced-form model when interpreting these results. These patterns show that an understanding of the relationship between personality and self-employment requires consideration of the impact of personality on both expected earnings and preferences, which cannot be decomposed from the results in Table 3. The structural model specified in the following section is designed to consider separately how personality affects the decision to become self-employed, both through its direct impact on preferences and indirectly through the effect on expected sectoral earnings.

²¹The standard errors in Column 1 are influenced by the skewness of the self-employment earnings distribution, in particular the presence of outliers, including earnings more than \$750,000. In results available from the authors, we regress earnings in self- or paid employment on the same set of variables as in Table 3, including a full set of interactions for self-employment. Once we trim outliers, we find that the coefficient on self-employment interacted with openness is negative and statistically significant at conventional levels, which means that individuals with this personality trait earn less in self-employment compared to paid employment. Given how outliers can affect parameter estimates, the structural model introduced in Section 4 incorporates a mixture distribution for business ideas to accommodate the long right tail.

²²In Tables S2 and S3 (available in the appendix), we report estimates from a series of probit models, where the outcome variable is an indicator for self-employment in 2004. Results on openness are robust to a number of specifications. In results available upon request, we also show that results are robust if we limit attention to the 726 individuals in our sample who were not self-employed in 1995. Of these, 65 (8.95%) report self-employment in 2004.

4 Model

This section specifies a model of selection into self-employment that incorporates socio-emotional skill and can be used to evaluate the impact of potential policies, such as subsidies for small business owners. In the model, individuals choose the sector with the highest returns, composed of earnings and non-pecuniary benefits, the latter captured as sector-specific flow utility. Monetary returns to self-employment are a function of the individual's business idea along with an endogenous capital investment, which may be sub-optimally low due to limited asset holdings and credit constraints. Socio-emotional skills can have different impacts on earnings in each sector and on sector preferences.

Recall that we highlight two key features of the model which are important for policy conclusions. First, the model captures several sources of selection into self-employment: credit constraints, flow utility and relative earnings. This means there are multiple reasons why an individual with a high-quality business idea may not choose self-employment. Credit-constraints can limit investments leading to low returns, the individual may enjoy paid employment more than self-employment or paid-employment earnings could be higher.²³ A second feature of the model we highlight is that agents observe business ideas prior to entry, which provides them with information about the returns to their business. This stands in contrast to models assuming that individuals make sector choices based on mean monetary returns, which could vastly overstate what individuals expect to earn due to the skewed distribution of entrepreneurial earnings.

4.1 Business Ideas, Earnings and Non-Pecuniary Benefits

Agents indexed by i draw a business idea θ_i and then decide between paid and self-employment, choosing the option delivering the highest expected utility.²⁴ Total utility for sector s is denoted V^s , where $s \in \{\text{SE}, \text{PE}\}$ with SE and PE referring to self-employment and paid employment, respectively. Utility in sector s is composed of income I^s and flow utility \tilde{u}^s . Each of these will be derived below.

Entrepreneurial earnings are generated according to the production function

$$y_i = \theta_i k_i^\alpha \xi_i \tag{1}$$

²³The latter point is made in Prada and Urzúa (2016) regarding mechanical skill, which has high returns among those who do not attain a college degree, helping to explain why individuals with the skill may forgo college.

²⁴We ignore non-workers and, therefore, selection into employment, though extending our analysis to include the decision to become employed would be straightforward.

where k_i is agent i 's capital invested in the entrepreneurial venture and $\alpha \in [0, 1]$ is a technology parameter that captures returns to capital. Our model of credit constraints follows Evans and Jovanovic (1989). By entering the model multiplicatively, high draws from the business idea distribution lead to a higher total and a higher marginal product of capital. One consequence of this specification is that agents with low reported assets can profit from a good idea despite constraints on their credit, which is in line with patterns in the data suggesting that low assets do not appear to preclude high earnings in self-employment. Rewrite this equation in logs:

$$\ln(y_i) = \ln(\theta_i) + \alpha \ln(k_i) + e_i^y \quad (2)$$

where $e_i^y \equiv \ln(\xi_i)$ and ξ_i is a disturbance term that is not observed by the agent before he chooses a sector. The distribution of e_i^y is specified below.

The business idea is generated as follows:

$$\ln(\theta_i) = x_i^\theta \beta^\theta + \sum_{j=1}^J \kappa_j^\theta f_{ij} + e_i^\theta \quad (3)$$

where x_i^θ is a vector of observable characteristics, β^θ is a vector of coefficients, and f_{ij} is personality trait $j \in \{1, \dots, 5\}$ for individual i which is mapped by κ_j^θ to the log value of the business idea.²⁵ e_i^θ captures factors affecting the business idea which are not observed by the econometrician, but are observed by the individual prior to the entry decision and is assumed to follow a mixed-normal distribution to account for the possibility of skew in entrepreneurial earnings.²⁶ Formally,

$$e_i^\theta \sim [p^\theta N(\mu_{\theta,1}, \sigma_{\theta,1}^2) + (1 - p^\theta) N(\mu_{\theta,2}, \sigma_{\theta,2}^2)]. \quad (4)$$

Substituting for θ_i in equation (2), we obtain the following expression for self-employment earnings:

$$\ln(y_i) = x_i^\theta \beta^\theta + \sum_{j=1}^J \kappa_j^\theta f_{ij} + e_i^\theta + \alpha \ln(k_i) + e_i^y. \quad (5)$$

Whereas e_i^y is a post-decision disturbance, e_i^θ is observed by the agent prior to his decision and thus must be integrated out of the model. This distinguishes our model from one in which the agent does not know e_i^θ prior to entry and instead the expected value of θ_i helps

²⁵As will be explained in Section 5, the five f_{ij} are possibly correlated latent factors identified from two measures of each of the five personality traits for each individual.

²⁶Estimating means of e_i^θ implies that equation (3) does not include a constant.

to explain sector choices. k_i is likewise not observed by the econometrician, but is derived from the optimal capital investment for a given business idea θ_i subject to possible credit constraints.

If the agent chooses paid employment, he earns wage w_i :

$$\ln(w_i) = x_i^w \beta^w + \sum_{j=1}^J \kappa_j^w f_{ij} + e_i^w \quad (6)$$

where x_i^w is a row-vector of observable characteristics that influence wage with prices β^w , κ_j^w is the price of personality trait j in the wage sector and e_i^w is a disturbance term, the distribution of which is specified below.²⁷ Net income from self-employment and paid employment are given by

$$\begin{aligned} I_i^{SE} &= y_i + r(A_i - k_i) \\ I_i^{PE} &= w_i + rA_i \end{aligned} \quad (7)$$

respectively, where r is the risk-free interest rate and A_i are assets in 2003. If the individual is in paid employment, he earns w_i along with returns to his assets. If he chooses self-employment, he earns y_i along with returns on assets net of what is invested in his business ($A_i - k_i$).

A problem with our data is that we do not observe assets in 2003, but only in 1995 or 2004. Using 1995 assets as a proxy for 2003 assets introduces measurement error, while using 2004 assets could introduce bias if they reflect (rather than affect) 2004 self-employment decisions. We therefore approximate 2003 assets as follows:

$$\ln(A_i) = \ln(\tilde{A}_i) \beta^a + e_i^a \quad (8)$$

where \tilde{A}_i are observed 1995 assets, β_a measures asset growth until 2003 and e_i^a a disturbance term capturing variance in asset growth that is normally distributed with mean zero and variance σ_a^2 . We jointly estimate these parameters with other model parameters.²⁸ Credit

²⁷We model self-employment in greater detail than we model paid employment to account for the role of optimal investments based on a business idea and credit constraints. Moreover, we permit flexibility in the portion of the business idea that is unobserved to the researcher (but is assumed to be observed to the agent prior to entry) to capture the skewed distribution of self-employment earnings.

²⁸Identification is discussed in the following section. Modeling 2003 assets in this manner essentially treats 1995 assets as a “noisy” measure of 2003 assets. If we instead use 1995 assets as a proxy, most qualitative results remain unchanged, though the credit constraint is more noisily estimated and has a smaller impact on earnings. Another option is to estimate β^a and σ^a outside of the model by regressing 2004 assets onto 1995 assets and then predict 2003 assets. Doing so does introduces potential bias, but in practice yields asset growth predictions in line with estimated parameters. By estimating these parameters jointly, however, we can allow asset growth to be correlated with other unobservables, which we explain below when discussing the error structure and correlation across equations.

constraints are imposed upon the entrepreneur such that $k_i \leq \lambda A_i$, where $\lambda \geq 1$. The entrepreneur is a net borrower when $A_i < k_i^*$ and a net-saver when $A_i \geq k_i^*$, where k_i^* denotes the optimal investment.²⁹

The agent chooses the sector $s \in \{\text{SE}, \text{PE}\}$ that generates the highest expected utility V_i^s given by

$$V_i^s = \rho I_i^s + \tilde{u}_i^s, \quad (9)$$

where \tilde{u}_i^s are non-pecuniary returns for sector s and ρ is a scaling parameter that converts dollars to utils. As we can only identify differences in non-pecuniary returns from choosing one sector versus the other, we specify net non-pecuniary benefits to self-employment as:

$$u_i^{SE} = \tilde{u}_i^{SE} - \tilde{u}_i^{PE} \equiv z_i \gamma^{SE} \quad (10)$$

which is equivalent to setting $\tilde{u}^{PE} = 0$. Here, z_i is a vector of characteristics and γ^{SE} are net non-pecuniary returns to observable characteristics in self-employment. z_i contains observable variables that are not included in the returns equations, such as spouse education, spousal employment and number of children.

4.2 Correlation Across Earnings Equations

Recall that while the agent observes e_i^θ prior to entry, e_i^y and e_i^w are post-entry earnings shocks in self- and paid employment, respectively. We permit correlation across the earnings equations by assuming the following error structure.

$$\begin{aligned} e_i^y &= \delta_1^y e_i^\theta + \delta_2^y \tilde{A}_i + \delta_3^y e_i^a + \nu_i^y \\ e_i^w &= \delta_1^w e_i^\theta + \delta_2^w \tilde{A}_i + \delta_3^w e_i^a + \nu_i^w \end{aligned} \quad (11)$$

where

$$\begin{pmatrix} \nu_i^y \\ \nu_i^w \end{pmatrix} \sim N \left[\begin{pmatrix} -\sigma_{y,\nu}^2/2 \\ -\sigma_{w,\nu}^2/2 \end{pmatrix}, \begin{pmatrix} \sigma_{y,\nu}^2 & 0 \\ 0 & \sigma_{w,\nu}^2 \end{pmatrix} \right] \quad (12)$$

The error structure means that the post-entry sector-specific earnings disturbances e_i^y and e_i^w are not assumed independent. Rather, they are dependent via realizations observed by the agent prior to entry, including 1995 assets \tilde{A}_i , the asset growth shock e_i^a and the portion of the business idea unobserved to the econometrician e_i^θ . Intuitively, this structure

²⁹The credit constraint is homogenous across groups, which could be relaxed. However, when we interact λ with other variables, we fail to detect significant differences.

permits 1995 assets, along with asset growth between 1995 and 2003 and the business idea draw, to provide information to the agent about factors affecting post-entry earnings shocks.

4.3 Optimal Investment and Expected Earnings

When deciding between sectors, the agent first computes expected self-employment earnings. He computes the optimal choice of k_i (supposing θ_i is known) by solving the following maximization problem:

$$\begin{aligned}
\max_k \quad & E[V_i^{SE}|\theta_i, A_i] \\
= \quad & E[I_i^{SE} + u_i^{SE}|\theta_i, A_i] \\
= \quad & E[\rho(y_i + rA_i - rk_i) + u_i^{SE}|\theta_i, A_i] \\
= \quad & E[\rho\theta_i k_i^\alpha \xi_i + \rho rA_i - \rho rk_i + u_i^{SE}|\theta_i, A_i] \\
= \quad & \theta_i k_i^\alpha E[\xi_i|\theta_i, A_i] - rk_i.
\end{aligned} \tag{13}$$

ξ_i is equal to $\exp(e_i^y)$, which is log-normally distributed.³⁰ Its expectation is thus given by:

$$\begin{aligned}
E[\exp(e_i^y)] &= \exp[\delta_1^y e_i^\theta + \delta_2^y \tilde{A}_i + \delta_3^y e_i^a + \sigma_{y,\nu}^2/2 - \sigma_{y,\nu}^2/2] \\
&= \exp[\delta_1^y e_i^\theta + \delta_2^y \tilde{A}_i + \delta_3^y e_i^a]
\end{aligned} \tag{14}$$

The agent's maximization problem conditional on entry is thus:

$$\begin{aligned}
\max_k \quad & E[V_i^{SE}] \\
= \quad & \theta_i k_i^\alpha [\exp[\delta_1^y e_i^\theta + \delta_2^y \tilde{A}_i + \delta_3^y e_i^a]] - rk_i \\
\equiv \quad & \theta_i k_i^\alpha \psi_i - rk_i.
\end{aligned} \tag{15}$$

where we set $\psi_i = \exp[\delta_1^y e_i^\theta + \delta_2^y \tilde{A}_i + \delta_3^y e_i^a]$.³¹ This means that:

$$k_i^* = \left(\frac{\alpha \psi_i \theta_i}{r} \right)^{\frac{1}{1-\alpha}}, \tag{16}$$

Plugging optimal capital k_i^* into the credit constraint inequality implies that entrepreneur is credit-constrained whenever:

$$\theta_i \psi_i > \frac{r}{\alpha} (\lambda A_i)^{1-\alpha}. \tag{17}$$

³⁰If we instead assumed independence of e_i^y and e_i^w , we would obtain that $E[\xi_i|e_i^\theta, A_i] = E[\xi_i] = 1$. In results available from the authors, we show that results using this simplifying assumption are qualitatively similar to our findings.

³¹Note that if the δ^y are all equal to zero (or if we assume that e_i^y and e_i^w are independent), then $\psi_i = 1$.

The inequality means that the agent is credit-constrained when he has a very good business idea but low assets. Expected paid-employment earnings prior to the sector choice also depend on A_i , e^a and e_i^θ and are given by:

$$\begin{aligned}\bar{w}_i &= \text{E}[w_i|e_i^\theta, A_i] = x_i^w \beta^w + \text{E}[e_i^w|e_i^\theta, A_i] \\ &= x_i^w \beta^w + \delta_1^w e_i^\theta + \delta_2^w \tilde{A}_i + \delta_3^w e_i^a - \sigma_{w,\nu}^2/2\end{aligned}\tag{18}$$

The value of the optimal sector choice, denoted V_i^* , is given by:

$$V_i^* = \begin{cases} \max\{((\phi^\alpha - r\phi)(\theta_i\psi_i)^{\frac{1}{1-\alpha}} + \frac{1}{\rho}u_i^{SE}, \bar{w}_i)\} & \text{if } \theta_i\psi_i \leq \frac{r}{\alpha}(\lambda A_i)^{1-\alpha} \\ \max\{(\theta_i(\lambda A_i)^\alpha - r\lambda A_i + \frac{1}{\rho}u_i^{SE}, \bar{w}_i)\} & \text{if } \theta_i\psi_i > \frac{r}{\alpha}(\lambda A_i)^{1-\alpha} \end{cases}\tag{19}$$

where ψ_i is defined as above and where $\phi_i \equiv (\frac{\alpha}{r})^{\frac{1}{1-\alpha}}$.

5 Estimation

Given the specification of the model, the vector of parameters to be estimated is:

$$\Phi \equiv [\alpha, \beta^\theta, \kappa^\theta, p^\theta, \mu_{\theta,1}, \mu_{\theta,2}, \sigma_{\theta,1}^2, \sigma_{\theta,2}^2, \beta^w, \kappa^w, \beta^a, \sigma_a^2, \rho, \delta^y, \delta^w, \sigma_{y,\nu}^2, \sigma_{w,\nu}^2].$$

This section describes identification and estimation of Φ . We also discuss how we incorporate socio-emotional skill, using multiple scores from personality tests to identify latent factors purged of measurement error.

5.1 Identification of Model Parameters

For each individual, conditional on expected earnings differences across sectors, utility parameters are identified from sector choices. The credit constraint parameter λ is identified from covariation in assets, self-employment entry and earnings. For example, if individuals with similar observable characteristics, but different assets, exhibit similar entry decisions and self-employment earnings patterns, this would suggest that credit constraints do not play an important role, leading to higher estimates of λ .

An important threat to identification arises because earnings parameters in each sector are estimated from individuals sorting into that sector, potentially introducing selection bias. To avoid bias, the model explicitly accounts for selection across sectors through earnings differences, preferences and assets, which can affect investments in business ideas and thus selection into self-employment. Identification relies on the assumption that our specifications

for the earnings processes, along with our incorporation of various selection mechanisms, are rich enough to allow us to use earnings parameters estimated from individuals in one sector to calculate expected earnings for individuals who are observed in the other sector.

Regarding the earnings processes, the key identifying assumption is that once we account for selection through credit constraints and preferences and have conditioned on observable characteristics, including personality traits, e_i^y and e_i^w (shocks to self-employment and paid-employment earnings, respectively) are dependent through initial assets, assets growth and the unobserved portion of the idea e_i^{θ} . Otherwise, they are independent. If not, i.e., if there are further omitted factors, the concern is that individuals sort into sectors because they expect, for instance, higher ν_i^y or lower ν_i^w .

Preferences are a key source of selection into employment sectors. Omitting them could induce selection bias in estimated earnings parameters. For example, strong preferences for self-employment would lead to entry despite relatively low expected earnings. Failing to account for preferences in modeling sector choices implies that earnings estimates would be upwardly biased to rationalize observed entry decisions. To help to separately identify preferences versus returns, we include variables in the utility function that shift preferences over sectors, but which are excluded from the earnings function. These include: number of children, spouse education and spouse employment. The assumption is that, conditional on variables included in the earnings functions (including marriage) excluded variables do not directly affect sector-specific pecuniary returns, but do affect sector preferences. These exclusion restrictions may be problematic if, for example, children help out in the family business, thus raising earnings. We cannot rule out this possibility, though argue that the first-order effect of variables capturing family structure is on selection into self-employment versus earnings. See Rees and Shah (1986) for an earlier discussion on family variables and selection into self-employment.³²

Asset holdings are another factor affecting selection into sectors and we incorporate this possibility into the model. For a given business idea and expected paid-employment earnings, individuals with higher asset holdings are able to invest more in high-quality businesses, which affects the entry decision. A problem incorporating assets is that they are potentially endogenous to the earnings shocks. For example, individuals who tend to generate high quality business ideas may accumulate more assets over time. Failing to account for this

³²Other possible exclusions include parents' education and variables indicating whether parents owned their own business. Not surprisingly, fathers' self-employment status is predictive of self-employment. However, we choose to omit it from our structural analysis. The reason is that personality may have a genetic component and so it is possible that father's self-employment captures an individual's personality. However, reduced-form choice models suggest that main results would not change if we added these variables. See results from reduced-form choice models (Tables S2 and S3) in the appendix.

possibility could bias estimates of credit constraints. For example, suppose individuals who expect to earn less in paid employment due to unobservable factors have fewer assets to invest in their business. If we fail to account for this correlation, higher entry due to low paid-employment earnings would suggest that assets are not very important, leading us to understate the importance of credit constraints. To address this potential source of bias, as discussed, earnings shocks are a function of 1995 asset holdings and the asset shock e_i^a . Earnings shocks are thus specified to capture the idea that unobserved factors affecting earnings across sectors could include factors related to 1995 assets or asset growth.

5.2 Latent Factors and Measurement Error

A key advantage of our approach is the incorporation of the Big 5 personality traits into a model of selection into self-employment. However, assessments of socio-emotional skills are subject to measurement error. In particular, individuals with a given underlying personality trait could score differently on tests that measure personality, depending on their observed characteristics, such as their age. This raises additional concerns about identification since estimates of parameters governing how personality affects earnings or utility would capture the impact of these other characteristics.

Fortunately, for each personality trait, the MIDUS data set includes two assessment scores for each individual. We use the two assessments to identify five latent personality traits f_{ij} , which enter earnings and the utility functions and which are purged of measurement error. This approach reduces standard errors, improving inference. It also has the advantage of providing a natural way to incorporate multiple assessments into the analysis.³³

Personality test scores are modeled as a function of latent factors, along with observable characteristics and random measurement error. For $J = 5$ latent factors, an observed measurement of skill $j \in \{1, \dots, J\}$ for person i at time t is denoted C_{ijt} and specified as:

$$C_{ijt} = M_{it}\rho_{jt} + d_{jt}^C f_{ij} + \epsilon_{ijt}^C \quad (20)$$

where M_{it} is a row-vector of observed characteristics with the accompanying vector of coefficients ρ_{jt} , f_{ij} is the value of latent skill j for person i , d_{jt}^C is the period- t factor loading

³³A related concern in linking personality to self-employment is reverse causality, the idea that self-employment decisions might affect personality. We cannot directly test this hypothesis using latent factors. However, we can test whether there is evidence for this type of reverse causality by examining personality assessment scores for individuals over time. In Table S4 in the appendix, we show that personality scores exhibit strong within-individual persistence over time. However, conditional on 1995 assessments for each of the five traits, 2004 assessments are not related to 1995 self-employment, 1995 assets or 1995 earnings.

on trait j and ϵ_{ijt}^C is an error term capturing mis-measurement.³⁴ M_{it} includes age and education at time t as well as the individual’s fluid cognitive ability test. Including these variables allows them to partially explain measurement error, which is the difference between the latent factor and the personality assessment. To fix ideas, conditional on a latent personality trait f_{ij} , an individual may exhibit a different score on the personality assessment C_{itj} due to age, education or cognitive ability. The latent and stable personality traits that we recover and which enter into the utility and earnings functions are therefore purged of average differences across the sample that are attributed to these observable factors. Latent factors f_{ij} are drawn from a multivariate joint distribution.³⁵

To achieve identification, a number of assumptions are required. First, we assume that $\text{cov}(f_{ij}, \epsilon_{ijt}^C) = 0 \forall t$ (latent traits are independent of measurement error) and that latent trait j does not affect the measured value of trait j' : $\text{cov}(C_{ij't}, f_{ijt}) = 0$ for $j \neq j', \forall t$. The latent traits are also assumed independent of the observables (age and cognition); i.e., age and cognition are treated as observable components of the measurement error affecting personality scores, but are not related to the underlying personality traits. Identification of parameters of the measurement system also requires normalizations. In particular, we set the first factor loading for each personality trait to 1. Moreover, we assume dedicated measurements, which means that functions of measurements for each trait (e.g., extraversion) only contain the corresponding underlying factor, i.e., we do not allow the factor “agreeableness” to explain the measurement of extraversion. Further details along with a proof of identification are in the Appendix.

Parameters of the measurement system are jointly estimated with all model parameters, which is reflected in the likelihood function derived in the Appendix.³⁶ Moreover, we augment the set of parameters due to joint estimation, which now includes Ξ_f , where Ξ_f includes all

³⁴Here, t refers to calendar time and is used to distinguish data collected in different years: 1995 and 2004.

³⁵In particular,

$$\begin{pmatrix} f_{iO} \\ f_{iC} \\ f_{iE} \\ f_{iA} \\ f_{iN} \end{pmatrix} \sim N \left[\begin{pmatrix} \mu_{iO} \\ \mu_{iC} \\ \mu_{iE} \\ \mu_{iA} \\ \mu_{iN} \end{pmatrix}, \begin{pmatrix} \sigma_O & \sigma_{O,C} & \sigma_{O,E} & \sigma_{O,A} & \sigma_{O,N} \\ \sigma_{O,C} & \sigma_C & \sigma_{C,E} & \sigma_{C,A} & \sigma_{C,N} \\ \sigma_{O,E} & \sigma_{C,E} & \sigma_E & \sigma_{E,A} & \sigma_{E,N} \\ \sigma_{O,A} & \sigma_{C,A} & \sigma_{E,A} & \sigma_A & \sigma_{A,N} \\ \sigma_{O,N} & \sigma_{C,N} & \sigma_{E,N} & \sigma_{A,N} & \sigma_N \end{pmatrix} \right] \quad (21)$$

³⁶We have experimented with alternative approaches, including assuming independence across factors (which also permits estimation of the measurement system in a separate first step), introducing more flexible function forms for factors (e.g., mixed normals) and omitting age and cognition from the measurement system. In general, our key qualitative results are not sensitive to these different assumptions on the measurement system, though some magnitudes change. Results are also robust to simply ignoring measurement error and using 1995 or 2004 measurements of personality, though standard errors are larger, which is expected.

parameters of the measurement system of the latent factors f_{ij} :

$$\Xi_f \equiv [\rho_{jt}, d_{jt}^C, \mu_j^C, \sigma_{j,j'}^C] \quad , \quad j, j' \in \{1, \dots, 5\}, t \in \{1995, 2004\}$$

5.3 Estimation Procedure Using Simulated Maximum Likelihood

We estimate the parameters of the model described in the previous section via simulated maximum likelihood. There are three main steps to the estimation procedure. First, at each set of parameter value suggestions, indexed by g and denoted $\Phi^{(g)}$, and for each individual i , we simulate earnings, personality traits and sector choice K times, where K represents the number of draws of unobservables for each individual.³⁷ Second, we compute each individual’s average likelihood contribution, where the average is taken over the K draws. Third, we sum over average likelihood contributions from each individual and compute the log, which yields the value of the simulated log likelihood function, the negative of which is then maximized as with standard likelihood functions. In this procedure, the business idea is “integrated out” in that it is assumed observed by each simulated agent prior to the entry decision and entry decisions are then averaged across simulated agents for each individual in the sample. Further details on the algorithm and the likelihood function are found in the appendix.

6 Parameter Estimates and Earnings Distributions

6.1 Sector-Specific Earnings Parameters

Earnings equations estimates are found in the first two columns of Panel A of Table 4. Beginning with personality traits, a contrast emerges regarding returns to the trait “openness to new experiences”. Though marginally profitable in paid employment, it lowers the quality of business ideas and therefore entrepreneurial earnings. Earlier work shows mixed results on returns to openness. Barrick and Mount (1991) show that open individuals are eager trainees though not better employees and Barrick and Mount (1993) find no evidence that open individuals fare better in jobs with greater autonomy. In some contexts, openness has been shown to predict a lack of commitment to an organization and a higher willingness to leave for a better opportunity (Moss et al., 2007). Openness may undermine business success if it means that individuals are less committed to a new business venture once it has

³⁷During estimation, we set $K = 2,500$.

been started.³⁸ In such cases, a more rigid focus on the new business may lead to greater success compared to imaginative thinking. We return to these points in the following section, when we discuss how openness to new experiences, despite being relatively unproductive in self-employment, also predicts a strong preference for it.

The second trait we examine, conscientiousness, is profitable in paid employment, though costly in self-employment. This latter finding is somewhat surprising as one would expect characteristics such as an attention to detail to be helpful in running a successful business. However, lacking conscientiousness is also related to self-indulgence or a tendency for ignoring rules for personal gain. This possibly links conscientiousness to a high disutility from breaking rules even if doing so will improve business performance. Consistent with this finding, Levine and Rubinstein (2017) find that deviant behaviors can be profitable in entrepreneurship. Further, a literature in personnel and organizational psychology has studied pro-social rule-breaking, also known as “constructive deviance,” whereby individuals break rules when it makes a business run better (Dahling et al., 2012). This research and our findings suggest that conscientious people could earn less in self-employment since they are inflexible or overly concerned with following rules even when doing so harms their business.

Extraversion is profitable in self-employment, which is consistent with previous work on personality and the labor market (Bowles, Gintis, and Osborne, 2001; Viinikainen et al., 2010; Caliendo, Fossen, and Kritikos, 2014). The impact of extraversion on paid-employment earnings is smaller and marginally significant. In contrast, agreeableness carries an earnings penalty in both sectors, though more strongly so in self-employment. In fact, earnings penalties for agreeableness have been shown in several studies (Heineck, 2010; Nyhus and Pons, 2005; Mueller and Plug, 2006). A key component of agreeableness is a lack of selfish behavior. Laboratory evidence has confirmed this. Ben-Ner, Kong, and Putterman (2004) relate behavior in dictator games to measurements of personality and find that agreeable individuals who are assigned the role of the dictator are more likely to offer higher amounts of money. Our results suggest that agreeableness may capture other-regarding or social preferences, altruism or a high psychic cost of profit-seeking at the expense of others.³⁹ It should therefore not be surprising that a trait capturing social preferences would carry an earnings penalty.

Neuroticism is profitable in both sectors, though the impact in self-employment is noisier, insignificant and smaller. Previous results on neuroticism are similarly mixed. Whereas Mueller and Plug (2006) and Heineck (2010) find a negative impact of neuroticism on earnings, Viinikainen et al. (2010) do not once they have controlled for work experience, which

³⁸A bit more bluntly, Oldham, Skodol, and Bender (2009) suggest that openness may reflect “flakiness”.

³⁹For a relatively early contribution on social preferences, see Kahneman, Knetsch, and Thaler (1986).

they offer as evidence that neuroticism leads to a less stable work history. This thinking is in line with the idea that neuroticism is linked to depression, which like other chronic illnesses can lead to gaps in work history (Artazcoz et al., 2004).

Turning to remaining earnings parameters, we find that in self-employment, individuals earn more when they are more highly educated and younger. Marriage also has a positive impact on earnings, but the parameter is insignificant. Fluid cognitive ability has a negative impact on earnings, which is not the expected sign. One possible explanation is that fluid cognitive ability is a mis-measurement of true cognition. Previous research has shown that the measure peaks around age 30 and declines thereafter.⁴⁰ Hence, the coefficient may capture avenues through which aging encourages less productive entrepreneurs to start a business. Assets tend to rise with age, in which case the coefficient could reflect how older agents face relaxed credit constraints, which could encourage entry with less lucrative ideas.⁴¹ In paid employment, education has a positive effect, which is weaker than its impact on self-employment earnings.⁴² Moreover, age, marriage and cognition lead to higher earnings in paid employment.

6.2 Preference Parameters

Utility parameters are found in the third column of Panel A in Table 4. These parameters capture relative utility in self-employment versus paid employment and can help to explain why individuals choose sectors in which they expect to earn relatively little. We find that openness to new experiences, though it generates relatively low-quality business ideas, captures a strong preference for running a business. Openness to new experiences is characterized by an attraction to new ideas and novel experiences. A new business venture might therefore be enticing for individuals with this personality trait. However, when coupled with earlier results about performance in self-employment, this finding suggests that the relationship between openness and business ownership is more nuanced. Open individuals may perform poorly once the novelty of a new venture wears off and the drudgery of running a business sets in. As discussed earlier, Moss et al. (2007) report an inverse relationship between openness and commitment to an organization. They also find that this negative

⁴⁰See, for example Horn and Cattell (1967) and Bugg et al. (2006).

⁴¹The negative coefficient is also present in the reduced-form earnings regressions; see Table 3. It is also worth noting that fluid cognitive ability is only measured once in the MIDUS data, which means we do not have enough data to separately identify a latent cognitive factor separately from a mis-measurement that we can allow to be a function of age to explicitly test this hypothesis. These types of problems underscore the value of applying methods that isolate latent, potentially mis-measured factors, which is what we do in the case of personality traits.

⁴²Our finding that education has high returns in self-employment in the U.S. accords with results from a meta-analysis reported in Van der Sluis, Van Praag, and Vijverberg (2008).

relationship is stronger when organizational resources are limited, which likely characterizes new businesses. Remaining estimates suggest that conscientiousness, extraversion and agreeableness capture preferences for paid employment, while neuroticism captures a preference for self-employment.

Results on openness and extraversion illustrate the added value of our approach. Both predict self-employment, but for different reasons. Open individuals start businesses because they enjoy it and not because they are good at it. Extraverted individuals do so because they generate high-value ideas. These distinctions would be difficult to identify absent the type of choice model we estimate. Moreover, this distinction will have implications for the impact of policies.

Turning to other sociodemographic variables, we find that educated individuals face lower relative utility in self-employment, which may arise if highly educated individuals who are paid employees have some of the benefits of self-employment, such as flexibility or autonomy. Age and marriage also predict higher relative utility as an entrepreneur, though the parameter for marriage is not significant. Fluid cognitive ability predicts a stronger preference for entrepreneurship. This finding suggests that intelligent people are more likely to be entrepreneurs, but parameter estimates in the earnings equations suggest they may not be particularly successful in terms of earnings.

The next parameters in Column [3] are for variables excluded from the earnings equation, but which play a role in the decision to enter self-employment. For example, spousal employment in 1995, which is a noisy measure of spousal employment in 2004, induces men to choose self-employment. This may be a signal that self-employment entails a lower cost in families with a second, steady income. Having more children seems to lower the desire to enter self-employment, but this parameter is insignificant. The second-to-last parameter in the panel converts utils to dollars.

One concern is that we do not model dynamics, so that preference parameters could capture future earnings. For example, among individuals who are open to new experiences, what we interpret as a strong taste for self-employment might instead capture expectations over future success in business, which is omitted from the model. Using data from 1995, we find little evidence that this is the case. Among individuals who are entrepreneurs in 1995, we would be concerned if openness predicted continued self-employment in 2004. Instead, we find that 2004 self-employment is negatively related to openness among individuals who were self-employed in 1995, though the parameter is insignificant. Thus, there is no evidence that openness leads to longer survival in self-employment. We would also be concerned that utility captures differences in earnings profiles if, among individuals who are self-employed in 1995

and in 2004, we found that openness predicted higher earnings in 2004. This would suggest that open individuals start out earning little as entrepreneurs, but foresee higher future returns, which our static model captures as utility. However, we find that openness has no relationship to 2004 earnings in self-employment for individuals who are self-employed in 1995.⁴³ In general, these findings suggests that our results on personality, preferences and sector choices are not driven by first-year earnings being un-representative of future entrepreneurial success.⁴⁴

6.3 Additional Structural Parameters

6.3.1 Business Ideas, Technology, Credit Constraints and Assets

Panel B of Table 4 presents estimates of additional structural parameters. Column [1] contains parameters describing the mixed normal distribution of business ideas. The long right tail of the distribution of self-employment earnings is captured by a relatively low probability (21%) of a high mean draw. The first estimate in Column [2] of Panel [B] is the technology parameter, α , which maps business ideas into earnings. It is estimated at 0.14 and is statistically significant. The credit constraint parameter λ is 1.94, which means that individuals can invest roughly twice their reported assets in a business venture (Evans and Jovanovic, 1989). The estimate is noisy, with a standard error of 1.67. This is not surprising given the weak empirical relationship between 1995 assets and self-employment earnings in 2004. The finding also foreshadows results from policy simulations showing that relaxing credit constraints does relatively little to change which ideas enter the market, which is a key finding of our study. In other words, although we cannot reject less binding credit constraints, even if we take the low point estimate at face value, relaxing credit constraints does very little to affect entry.

The remaining two parameters in Column [2] of Panel C suggest that unobserved 2003 assets A_i (which enter the decision problem) are expected to be slightly higher than 1995 assets and that the error term has a large variance (4.60).⁴⁵

⁴³These findings are presented in Tables S7 and S8 in the appendix.

⁴⁴Some recent papers on self-employment model dynamics (Humphries, 2017; Dillon and Stanton, 2018; Hincapié, 2018). While these papers can focus on aspects of self-employment that we do not (e.g., learning), our model handles both information about business ideas and credit constraints more formally, which means our model is well-suited for the types of counterfactuals we perform below.

⁴⁵Similar parameter estimates are obtained by regressing observed 2004 assets onto 1995 assets, even though we do not explicitly match 2004 assets in our estimation procedure.

6.3.2 Earnings Shocks

Panel C of Table 4 presents estimates of parameters governing the distribution of post-entry earnings shocks e^w and e^y . These shocks are correlated through joint correlation with the realized business idea e^θ , starting assets \tilde{A} and the asset shock e^a . We find that higher starting assets, higher asset growth and better business ideas are all associated with higher earnings shocks in both sectors. In other words, assets accumulated by 2003 are endogenous to unobserved factors affecting expected sector-specific earnings. Intuitively, individuals with high-quality ideas and high levels of accumulated assets by 2003 (through higher 1995 assets or faster asset growth between 1995 and 2003) also expect relatively high post-entry earnings shocks in both employment sectors.

6.3.3 Measurement of Latent Traits

Estimated coefficients of the measurement system that relates latent, stable personality traits to scores from personality assessments are presented in Tables S5 and S6. We begin with a discussion of correlation among factors, coefficients for which are presented in the first part of Table S5. We find that, in the population, underlying personality traits are correlated. Openness is positively correlated with conscientiousness, extraversion and agreeableness and is negatively correlated with neuroticism. Conscientious individuals tend to be more extraverted and agreeable, but less neurotic. Extraversion and agreeableness exhibit a strong positive correlation. Finally neuroticism is negatively correlated with all other traits. It is important to account for these correlations. For example, ignoring the correlation could lead to over-estimation of the utility open individuals gain from self-employment if we ignore that open individuals tend to be extraverted, which raises expected self-employment earnings through better business ideas.

Means for each personality trait are not very far from raw data means of the personality assessments though variance is significant, implying that measurement error could be a concern if we simply included both 1995 and 2004 measurements (or some combination of the two) in our earnings and utility equations. Factor loadings tend to be near one. In general, the education parameters are negative, small and often insignificant. The interpretation is that, for a given latent personality trait, individuals who are more highly educated exhibit lower measures of each trait on the assessments, the exception being openness. Cognition also exhibits weak and generally insignificant relationships with departures of personality assessments from the latent personality trait. Coefficients are positive for openness, extraversion and agreeableness, but negative for the other traits. Some interesting patterns emerge as individuals age. For example, the age parameter in the measurement equation

for openness is -0.012 in 1995 and -0.007 in 2004. This means that, for a given underlying openness factor, individuals would be assessed as less open as they age, though this effect is small in 2004 versus 1995. Another example is extraversion. In 1995, the age parameter is negative, but it is positive in 2004, which means that older individuals are assessed to be more extraverted than they actually are in 2004. Generally, parameters in Tables S5 and S6 suggest that measurement error exists, which means that inference is improved if we exploit multiple measures of personality to identify the distribution of latent traits, which are then used in the choice and earnings equations.

6.4 Model Fit and Earnings Distributions

We simulate earnings distributions to assess model fit and to illustrate the skewed distribution of self-employment earnings. In particular, for each of the 896 individuals in our sample, we simulate 2500 observations, which amounts to simulating roughly 2.2 million workers with the same distribution of observable characteristics as individuals in our sample. For each observation, we then draw business ideas, utility, expected earnings, sector choice and then sector-specific shocks. The predicted probability of self-employment is 19%, which corresponds to the observed probability from the data. We also plot histograms of sector specific earnings for individuals observed in each sector in Figures 2(a) and 2(b) where, for comparison, we also plot observed earnings. Notice that in both sectors earnings are considerably skewed, which is captured quite well by the estimated model.

Next, in Figure 3, we plot earnings differentials (self versus paid) for each tenth of a percentile. The x -axis is the percentile tenth and the y -axis is the expected earnings differential. Figure 3 shows that our model captures earnings patterns that are important for understanding entrepreneurship. First, simple comparison of average earnings across sectors can be misleading (Hamilton, 2000). Roughly 82% of draws plotted in Figure 3 are below zero. In other words, if each draw is interpreted as a simulated worker, 82% of simulated workers would earn less by choosing self-employment.⁴⁶ Second, there is a small probability of earning substantial returns in entrepreneurship. For example, a draw in the top 0.1% generates an earnings differential up to \$5,000,000. These exceedingly high, and exceedingly rare, draws are enough to drive up within-individual averages considerably. The figure essentially illustrates the observation that if Bill Gates walks into a bar, the average individual in the bar is a multi-millionaire. The typical individual, however, is not. Figure 3 therefore illustrates why a model using average earnings as an additional regressor in a choice model

⁴⁶Of the remaining 18% who would earn more, some would still opt for paid employment due to the opportunity cost of investing assets in their business rather than investing it elsewhere.

(e.g., a “structural probit”) to explain entry could be highly misleading. The skewed distribution means that using average earnings to compute expected earnings in self-employment could lead to an upward bias. To rationalize low entry probability, utility would thus be biased downwards. This would in turn suggest that there is a large group of “reluctant entrepreneurs,” those with high expected earnings, but low utility from running a business. If so, policies attracting individuals into entrepreneurship could appear to be more beneficial than they are.

6.5 Misaligned Preferences and Performance in Self Employment

A strength of our modeling approach we illustrate next, and which distinguishes this paper from earlier work linking socio-emotional skill to entrepreneurial behavior, is that it allows us to identify characteristics that predict strong preferences for sectors in which they are relatively unproductive. Here, we highlight estimates where preferences and productivity are misaligned, driving agents toward a sector in which they will earn less. Next, we examine the relationship between extraversion, openness and self-employment more closely since both predict entry, but — as parameter estimates suggest — for very different reasons.

Table 5 lists the characteristics affecting both earnings and utility, indicating those where a misalignment exists. According to Table 5, higher education is associated with a stronger preference for paid employment even though higher education leads to higher relative earnings in self-employment. Age works the opposite way. Older agents may prefer self-employment despite earning more in paid employment in part since self-employment offers greater flexibility as agents move towards semi-retirement.⁴⁷ Turning to personality traits, consider the distinction between openness and agreeableness. Both generate higher relative earnings in paid employment. However, while agreeable individuals prefer paid employment, individuals who are open to new experiences prefer self-employment. In other words, there is a misalignment because open individuals prefer a sector in which they are relatively unproductive.

6.6 Externalizing, Openness and Entry: An Illustration

To further highlight the value of our approach, we examine in more depth the various ways in which two personality traits, extraversion and openness, affect the entry decision. Previous literature has recognized their association with self-employment. A benefit of our approach is

⁴⁷Alternatively, older self-employed agents would not have to contend with the risk of being replaced by younger employees with lower tenure who are therefore cheaper to employ.

that it allows us to show that the mechanisms underlying these relationships are different. While extraverted individuals are attracted to self-employment because they generate high-quality business ideas, open individuals start businesses because they enjoy it.

To illustrate these mechanisms, we compute deciles of extraversion and openness. Then, for each of the resulting possible one hundred combinations, we set traits to these levels for each individual in the sample. Finally, we simulate optimal earnings and decisions to illustrate the labor market impact of latent, stable personality traits. Figure 4(a) plots earnings in self-employment for different combinations of values of extraversion and openness applied to all individuals in the sample regardless of their optimal choice. We see that low levels of extraversion combined with high levels of openness have the starkest income penalties in self-employment. Introverted individuals who are open to new experiences can expect the lowest returns to opening their own business. The differences are not small, with average earnings ranging between about \$22,000 and \$50,000. In paid employment, the highest wage penalties accrue to those who are neither extraverted nor open to new experiences (Figure 4(b)). Here, the range of earnings is smaller: about \$70,000 to \$84,000. Figure 4(c) plots utility u_i^{SE} (converted to dollars using the estimated multiplier). Utility of self-employment ranges from about \$10,000 to about \$40,000. This range helps to explain why many individuals choose to run a business that will not be particularly lucrative: the enjoyment of doing so is worth thousands of dollars.

Misalignment suggests that the relationship between personality traits and sorting into self-employment is complex and that average entry for different combinations of traits can also be non-monotonic. According to Figure 4(d), higher levels of extraversion encourage entry at low levels of openness. As openness increases, entry rises. At high levels of openness, extraversion can even reduce entry. The interpretation is that at low levels of openness, extraversion encourages entry through better business ideas, while at high levels of openness, extraversion discourages entry through its impact on utility.

7 Credit Constraints, Preferences and Subsidies

This section uses the estimated model to perform several experiments. The first set examines impediments to high-quality businesses by assessing the role of preferences and constraints in sector choices. Next, we consider a counterfactual subsidy of \$25,000 for business owners. For each experiment, we simulate earnings differentials and entry decisions for each individual in the sample 2500 times. Earnings differentials are one measure of business quality, but are also affected by capital investment decisions and credit constraints, which depend on agent

assets. Thus, to measure the impact of policies on business quality, rather than individual returns, we also examine changes to the value of business ideas θ_i , which provide an arguably better measure of the social benefits of businesses.⁴⁸

7.1 Credit Constraints, Preferences and Starting a Business

Table 6 provides results from counterfactuals, in particular removing credit constraints and the role of preferences. Recall that average entry probability predicted by the model is 19%, which aligns with observed entry probability. The second row shows that when we simulate the removal of liquidity constraints, self-employment entry increases to 22%. Moreover, average self-employment earnings rise slightly (by about 0.2%). This could be due to better ideas or higher investments in high quality ideas. By simulating the value of realized ideas, we show that the increase in earnings is due to the latter. In fact, the average value of realized ideas is slightly lower. This means that credit constraints, rather than obstructing the realization of good business ideas generated by agents with few assets, instead lead some firms to operate at sub-optimal scale (and perhaps also exert a screening effect, which is however very small).

In the third row of Table 6 we assume that agents choose the sector that maximizes earnings (liquidity constraints are imposed). The goal is to assess whether self-employment is driven by reluctant entrepreneurs with good ideas who would otherwise prefer paid employment or, alternatively, it is dominated by individuals who enjoy opening a business despite low relative returns. We find clear evidence for the latter. In the absence of the non-pecuniary benefits of entrepreneurship, self-employment entry declines to 13%, which reflects the reduction in entry of lower-quality business ideas. Earnings rise by 30% and the value of ideas rises by 19%. When we assume earnings maximization and remove liquidity constraints, the impact on entry is slightly reduced, but again we find that individuals deterred from entrepreneurship due to credit constraints tend to be less productive entrepreneurs. To further illustrate these patterns, in Figures 5(a) and 5(b) we show the distributions of simulated relative earnings and the value of business ideas, respectively, under the baseline model and the two counterfactuals. Average income from self-employment is slightly higher absent credit constraints due to larger investments in very high-quality businesses, but eliminating a role for preferences shifts both distributions markedly to the right.

Overall, the simulations show that credit constraints play a limited role in preventing lower-quality businesses from reaching the market. This is consistent with Hurst and Lusardi

⁴⁸Given that business ideas map into earnings according to the production function, it is not clear that the level of the value of θ_i is meaningful, so we only report changes resulting from each policy experiment.

(2004), who show that liquidity constraints may force some entrepreneurs to operate at sub-optimal scale, but do not appear to affect the distribution of realized ideas. The logic is that if potential business owners have a very valuable idea, they may not be able to invest in it at an optimal level due to credit constraints. However, they will still start the business. Yet, given the noisiness of the estimated credit constraint parameter, we interpret this result with caution.

To further illustrate the roles of utility and credit constraints in selection into self-employment, we plot openness and the value of business ideas in Figures 5(c) and 5(d). The solid line at the bottom of Figure 5(c) shows the unconditional average value of business ideas for different levels of openness. The dashed line conditions on entry and, because it sits well above the solid line, shows positive selection at entry across the distribution of openness. The degree of this selection, though, appears to decrease as openness increases, as the conditional and unconditional average values are closer together at higher levels of openness. Removing credit constraints has very little effect on the average value of business ideas. However, eliminating the role of preferences (the dot-dashed line) further increases positive selection, particularly at high levels of openness. That is, if individuals' preferences are not permitted to affect whether they enter self-employment, selection into self-employment by the value of business ideas operates more strongly and to a more similar degree across the distribution of openness. This last point is made clearer by Figure 5(d), in which we normalize average business idea value to 1 at the lowest decile of openness under each counterfactual. This facilitates comparison of proportional changes; the degree to which increased openness proportionally decreases the value of business ideas in each scenario is the slope of the appropriate line in Figure 5(d), with the solid line representing all business ideas. The slope is steeper if we simply condition on entry. Unconditionally, high-openness ideas are worth 70% of low-quality ideas. Conditional on entry, this number is 62%. The reason is that conditioning on entry allows utility to play a role, and greater openness induces entry by entrepreneurs with lower-value ideas than would otherwise enter.⁴⁹ Accordingly, if preferences play no role, agents care only about relative earnings, and the relative decline in average idea quality with openness is smaller than in the unconditional distribution (high-openness ideas are now 85% the value of low-openness ideas). This is because high-openness individuals with low-quality ideas are no longer induced into business by their preferences, which are nullified here.

⁴⁹Again, according to the dot-dashed line, removing credit constraints does little to change the relationship between openness and the value of realized ideas.

7.2 The Impact of Subsidies Promoting Entrepreneurship

Many policymakers have considered how best to design and implement policies that foster successful entrepreneurship.⁵⁰ The motivation appears to be the following: if good ideas do not make it to the market, then society loses out on innovations, tax revenue, better products and — assuming path-dependence — future innovation. However, one should be careful in designing policies affecting self-employment. It may not be worthwhile to foster entrepreneurship *per se*, but to encourage the realization of good business ideas. In this section, we use our estimated model to assess a subsidy.⁵¹

Before presenting results, we briefly discuss reasons why according to our model a subsidy could change the average quality of ideas entering the market. If individuals were modeled simply as maximizing income when choosing a sector, a subsidy could only attract ideas by inducing individuals who earn more in paid employment to start a business. By incorporating preferences, we allow for the possibility that a subsidy could attract talented “reluctant entrepreneurs” who would otherwise choose paid employment. If this occurs, then policies promoting entrepreneurship could be worthwhile because they attract individuals with high-quality business ideas but an aversion to opening their own business. On the other hand, we would not expect a subsidy to attract individuals with extremely high-quality ideas because they would likely enter without the subsidy even with preferences for paid employment. Moreover, recall that our model captures the skewed distribution of self-employment earnings. This means that, despite high average earnings, most individuals do not earn more in self-employment, which means that preferences play an important role in the decision. A subsidy could be counterproductive if it attracts “lifestyle entrepreneurs” who would enjoy owning a business, but who refrain from doing so because they have low-quality ideas. In summary, in our framework the impact of a subsidy on the average quality of ideas is conceptually ambiguous and thus an empirical question. A related possibility that is worth mentioning is that government policy could help highly productive entrepreneurs to overcome their aversion to risk. Though we do not measure risk aversion explicitly, it is not the case that our main results on whether subsidies are worthwhile are therefore incorrect. Risk aversion is a preference and therefore it is captured in our utility function. Indeed, Åstebro et al. (2014) argue that evidence on whether entrepreneurs have different risk preferences

⁵⁰See, for example, Lerner (2009) for a discussion of the pitfalls associated with public policies promoting entrepreneurship and venture capital.

⁵¹For example, in the U.S. city of St. Louis, there are multiple sources of funding for startups. While some grants are for existing and profitable firms, there are also “pre-revenue,” equity-free grants aimed at funding startups, generally ranging between \$5,000 and \$50,000. Often, these are in the form of a competition, which would suggest a tournament funding only the highest-quality ideas. The implicit assumption behind the straight subsidy we examine is that the supply of such grants is elastic.

than paid employees is mixed and inconclusive.

We assess the impact of a subsidy that pays entrepreneurs \$25,000 if they open their own business.⁵² We also assess the effect of the policy on the distribution of realized business ideas. We find that the subsidy encourages the realization of relatively low-quality ideas. Rather than attracting reluctant but talented entrepreneurs — a possibility that has been used to justify government support of small business — the subsidy attracts people with low-quality ideas who would enjoy owning their own business. Absent the subsidy, these individuals stay in paid employment due to low expected earnings in self-employment.

The last two lines of Table 6 illustrate the impact of a subsidy. Entry nearly doubles from 19% to 37%. Earnings and the value of business ideas fall by 46% and 41%, respectively. We also find that preferences exacerbate the negative impact of a subsidy on the distribution of ideas that enter the market. If agents are assumed to be income maximizers, entry falls by 1 percentage point, with earnings and business idea value rising by 9%. This suggests that a key mechanism of the subsidy is to attract agents with a preference for self-employment, but low expected returns, into entrepreneurship, where they start businesses based on low-quality ideas.

Shifts in the distributions of entry probabilities and the value of realized ideas induced by the subsidy are plotted in Figures 6(a) and 6(b). Here, these shifts are measured using estimated preference parameters and again where preferences for the non-pecuniary aspects of entrepreneurship are assumed to play no role. These figures illustrate that the shifts are much sharper due to preferences. In Figures 6(c) and 6(d), we repeat the exercise from Figures 5(c) and 5(d). As before, we find clear evidence of positive selection into self-employment across the distribution of openness. Moreover, a subsidy steepens the relative loss of value of ideas as openness increases since individuals with low-quality ideas, but high utility from owning a business, are induced into the sector. As illustrated by the dotted line, this effect is attenuated in the presence of a subsidy where preferences play no role.

The notion that agents with personalities predicting low-quality ideas are being induced into entrepreneurship is illustrated in Figure 7, where the distribution of each personality trait in self-employment is plotted according to model parameters and again assuming the counterfactual subsidy. The subsidy attracts individuals who are open and agreeable. Interestingly, it also induces extraverted individuals, who tend to have higher-quality ideas but also prefer paid employment, a fact which should mitigate the negative impact of the subsidy on average ideas in general. On balance, however, an across-the-board subsidy for

⁵²For each draw, once an agent has solved the maximization problem conditional on the realization of his business idea, \$25,000 is added to his self-employment earnings. The linear structure of consumption utility means that the subsidy does not affect the optimal capital investment.

entrepreneurship seems ill-advised as it simply encourages less productive ventures, having the most dramatic effect on individuals who have low-quality ideas, but who nonetheless reveal a strong preference for entrepreneurship.

8 Conclusion

We provide a framework which captures how multiple dimensions of non-cognitive skill (as measured by the Big 5 personality traits) affect selection into entrepreneurship. Our framework allows the same personality traits to have different effects on sector-specific earnings and preferences. We also take explicit account of credit constraints along with the skewed distribution of earnings, the latter implying that the typical potential entrepreneur does not expect to earn a lot, but might start a business if it is enjoyable. We find that in many cases the same personality trait can predict both strong preferences for and poor performance in self-employment. Using the estimated model, we also find evidence that credit constraints do not appear to prevent high quality ideas from entering the market, though they might lead to sub-optimally low investments. Moreover, rather than attracting talented, reluctant entrepreneurs, a subsidy instead attracts individuals with personalities that generate low-quality ideas. While our model does not account for dynamics, we show that low self-employment income in the first year is not associated with steeper income profiles, which suggests that first-year earnings are an adequate proxy for the pecuniary returns to self-employment given the types of policies we use the model to evaluate.

A typical justification for policies aimed at encouraging small business ownership, such as subsidies (financed largely through taxation), is that society as a whole would benefit from unrealized ideas. Such arguments may be an artifact of using mean earnings from a highly skewed distribution to characterize the median potential entrepreneur's entry choice. We show that a subsidy attracts low-quality ideas.

Our findings suggest that subsidies may therefore require a different justification, one that is not rooted in the idea that there is a mass of productive potential entrepreneurs with an aversion to self-employment. One potential alternative justification is that entrepreneurship has positive externalities. For example, small businesses provide jobs and could thus help low-income communities where large corporations find it unprofitable to locate or provide value as an amenity if consumers enjoy enterprises that are independently owned. In other words, from a social welfare perspective, individual earnings are but one factor to consider. Indeed, Åstebro et al. (2014) suggests that 90% of the benefits of breakthrough innovation go to society as a whole rather than to the individual inventor, their partners, or their financial

backers. Assessing whether pricing this externality justifies tax-funded subsidies is beyond the scope of our model, which focuses on individual earnings. Thus, a fruitful — but by no means simple — direction for future research would examine an expanded set of social returns to entrepreneurship, such as positive externalities or spillover effects, which are not captured by individual earnings.

References

- Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. 2011. “Personality Psychology and Economics.” *Handbook of the Economics of Education* 4 (1).
- Artazcoz, Lucía, Joan Benach, Carme Borrell, and Immaculada Cortès. 2004. “Unemployment and Mental Health: Understanding the Interactions among Gender, Family Roles, and Social Class.” *American Journal of Public Health* 94 (1):82–88.
- Asoni, Andrea. 2010. “What Drives Entrepreneurship?” Mimeo, University of Chicago.
- Åstebro, Thomas, Holger Herz, Ramana Nanda, and Roberto A Weber. 2014. “Seeking the Roots of Entrepreneurship: Insights from Behavioral Economics.” *The Journal of Economic Perspectives* 28 (3):49–69.
- Åstebro, Thomas and Peter Thompson. 2011. “Entrepreneurs, Jacks of All Trades or Hobos?” *Research Policy* 40 (5):637–649.
- Barón, Juan D and Deborah A Cobb-Clark. 2010. “Are Young People’s Educational Outcomes Linked to their Sense of Control?” IZA Working Paper.
- Barrick, Murray R and Michael K Mount. 1991. “The Big Five Personality Dimensions and Job Performance: A Meta-Analysis.” *Personnel Psychology* 44 (1):1–26.
- . 1993. “Autonomy as a Moderator of the Relationships Between the Big Five Personality Dimensions and Job Performance.” *Journal of Applied Psychology* 78 (1):111.
- Becker, Gary S. 1965. “A Theory of the Allocation of Time.” *Economic Journal* 75 (299):493–517.
- Ben-Ner, Avner, Fanmin Kong, and Louis Putterman. 2004. “Share and Share Alike? Gender-Pairing, Personality, and Cognitive Ability as Determinants of Giving.” *Journal of Economic Psychology* 25 (5):581–589.
- Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel. 2008. “The Economics and Psychology of Personality Traits.” *Journal of Human Resources* 43 (4):972–1059.
- Borghans, Lex, Bart HH Golsteyn, James Heckman, and John Eric Humphries. 2011. “Identification Problems in Personality Psychology.” *Personality and Individual Differences* 51 (3):315–320.
- Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. “Incentive-Enhancing Preferences: Personality, Behavior, and Earnings.” *The American Economic Review* 91 (2):155–158.
- Brandstätter, Hermann. 2011. “Personality Aspects of Entrepreneurship: A Look at Five Meta-Analyses.” *Personality and Individual Differences* 51 (3):222–230.
- Bugg, Julie M, Nancy A Zook, Edward L DeLosh, Deana B Davalos, and Hasker P Davis. 2006. “Age Differences in Fluid Intelligence: Contributions of General Slowing and Frontal Decline.” *Brain and Cognition* 62 (1):9–16.

- Caliendo, Marco, Frank Fossen, and Alexander S Kritikos. 2014. "Personality Characteristics and the Decisions to Become and Stay Self-Employed." *Small Business Economics* 42 (4):787–814.
- Caspi, Avshalom. 2000. "The Child is Father of the Man: Personality Continuities from Childhood to Adulthood." *Journal of Personality and Social Psychology* 78 (1):158.
- Cattan, Sarah. 2011. "The Role of Workers Traits in Explaining the Early Career Gender Wage Gap." Mimeo, University of Chicago.
- Cobb-Clark, Deborah A and Stefanie Schurer. 2012. "The Stability of Big-Five Personality Traits." *Economics Letters* 115 (1):11–15.
- Cubel, Maria, Ana Nuevo-Chiquero, Santiago Sanchez-Pages, and Marian Vidal-Fernandez. 2016. "Do Personality Traits Affect Productivity? Evidence from the Laboratory." *The Economic Journal* 126 (592):654–681.
- Cunha, Flavio, James J Heckman, and Susanne M Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3):883–931.
- Cutler, David M. and Adriana Lleras-Muney. 2010. "Understanding Differences in Health Behaviors by Education." *Journal of Health Economics* 29 (1):1–28.
- Dahling, Jason J, Samantha L Chau, David M Mayer, and Jane B Gregory. 2012. "Breaking Rules for the Right Reasons? An Investigation of Pro-Social Rule Breaking." *Journal of Organizational Behavior* 33 (1):21–42.
- Dillon, Eleanor W. and Christopher T. Stanton. 2018. "Self-Employment Dynamics and the Returns to Entrepreneurship." Mimeo, HBS.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2008. "Representative Trust and Reciprocity: Prevalence and Determinants." *Economic Inquiry* 46 (1):84–90.
- . 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review* 100 (3):1238–1260.
- Evans, David S and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *The Journal of Political Economy* 97 (4):808–827.
- Fairlie, Robert W and William Holleran. 2012. "Entrepreneurship Training, Risk Aversion and other Personality Traits: Evidence from a Random Experiment." *Journal of Economic Psychology* 33 (2):366–378.
- Fairlie, Robert W, Dean Karlan, and Jonathan Zinman. 2015. "Behind the GATE Experiment: Evidence on Effects of and Rationales for Subsidized Entrepreneurship Training." *American Economic Journal: Economic Policy* 7 (2):125–161.
- Fréchette, Guillaume R, Andrew Schotter, and Isabel Trevino. 2017. "Personality, Information Acquisition, and Choice under Uncertainty: An Experimental Study."
- Gensowski, Miriam, James Heckman, and Peter Savelyev. 2011. "The Effects of Education, Personality, and IQ on Earnings of High-Ability Men." Mimeo, University of Chicago.
- Goldberg, Lewis R. 1971. "Five Models of Clinical Judgment: An Empirical Comparison between Linear and Nonlinear Representations of the Human Inference Process." *Organizational Behavior and Human Performance* 6 (4):458–479.
- Goldberger, Arthur S. 1972. "Structural Equation Methods in the Social Sciences." *Econometrica* 40 (6):979–1001.

- Hamilton, B.H. 2000. "Does Entrepreneurship Pay? An Empirical Analysis of the Returns of Self-Employment." *Journal of Political Economy* 108 (3):604–631.
- Hartog, Joop, Mirjam Van Praag, and Justin Van Der Sluis. 2010. "If You Are So Smart, Why Aren't You an Entrepreneur? Returns to Cognitive and Social Ability: Entrepreneurs versus Employees." *Journal of Economics & Management Strategy* 19 (4):947–989.
- Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes." *The American Economic Review* 103 (6):1–35.
- Heckman, James J. 2012. "The Developmental Origins of Health." *Health Economics* 21 (1):24–29.
- Heckman, James J and Bo E Honore. 1990. "The Empirical Content of the Roy Model." *Econometrica* 58 (5):1121–1149.
- Heckman, James J and Tim Kautz. 2012. "Hard Evidence on Soft Skills." *Labour Economics* 19 (4):451–464.
- . 2013. "Fostering and Measuring Skills: Interventions that Improve Character and Cognition." NBER working paper.
- Heckman, James J and Paul A LaFontaine. 2010. "The American High School Graduation Rate: Trends and Levels." *The Review of Economics and Statistics* 92 (2):244–262.
- Heckman, James J and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *The American Economic Review* 91 (2):145–149.
- Heckman, James J, Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *The Journal of Labor Economics* 24 (3):411.
- Heineck, Guido. 2010. "Does It Pay to Be Nice-Personality and Earnings in the United Kingdom." *Industrial and Labor Relations Review* 64:1020–1047.
- Hincapié, Andrés. 2018. "Entrepreneurship over the Life Cycle: Where are the Young Entrepreneurs?" Mimeo, UNC-Chapel Hill.
- Hisrich, Robert, Janice Langan-Fox, and Sharon Grant. 2007. "Entrepreneurship Research and Practice: A Call to Action for Psychology." *American Psychologist* 62 (6):575.
- Horn, John L and Raymond B Cattell. 1967. "Age Differences in Fluid and Crystallized Intelligence." *Acta Psychologica* 26:107–129.
- Humphries, John Eric. 2017. "The Causes and Consequences of Self-Employment over the Life Cycle." Mimeo, Yale University.
- Hurst, Erik and Annamaria Lusardi. 2004. "Liquidity Constraints, Household Wealth, and Entrepreneurship." *Journal of Political Economy* 112 (2):319–347.
- Hurst, Erik and Benjamin Pugsley. 2014. "Wealth, Tastes, and Entrepreneurial Choice." Mimeo, University of Chicago.
- Hurst, Erik and Benjamin Wild Pugsley. 2011. "What Do Small Businesses Do?" *Brookings Papers on Economic Activity* 2011 (2):73–118.
- Jöreskog, Karl G and Arthur S Goldberger. 1975. "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable." *Journal of the American Statistical Association* 70 (351a):631–639.

- Kahneman, Daniel, Jack L Knetsch, and Richard Thaler. 1986. "Fairness as a Constraint on Profit Seeking: Entitlements in the Market." *The American Economic Review* :728–741.
- Lazear, Edward P. 2004. "Entrepreneurship." *Journal of Labor Economics* 23 (4):pp. 649–680.
- Lerner, Josh. 2009. *Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed—and What to Do about It*. Princeton University Press.
- Levine, Ross and Yona Rubinstein. 2017. "Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?" *Quarterly Journal of Economics* 132 (2):963–1018.
- Lundberg, Shelly. 2011. "Psychology and Family Economics." *Perspektiven der Wirtschaftspolitik* 12 (s1):66–81.
- . 2012. "Personality and Marital Surplus." *IZA Journal of Labor Economics* 1 (1):1–21.
- . 2013. "The College Type: Personality and Educational Inequality." *Journal of Labor Economics* 31 (3):421–441.
- Lundborg, Petter. 2013. "The Health Returns to Schooling—What Can We Learn from Twins?" *Journal of Population Economics* 26 (2):673–701.
- MIDMAC. 1999. "Methodology of the National Survey of Midlife Development in the United States (MIDUS)." The John D. and Catherine T. MacArthur Foundation.
- Moss, Simon A, Janis McFarland, Simon Ngu, and Anna Kijowska. 2007. "Maintaining an Open Mind to Closed Individuals: The Effect of Resource Availability and Leadership Style on the Association between Openness to Experience and Organizational Commitment." *Journal of Research in Personality* 41 (2):259–275.
- Mueller, Gerrit and Erik Plug. 2006. "Estimating the Effect of Personality on Male and Female Earnings." *Industrial and Labor Relations Review* :3–22.
- Nyhus, Ellen K and Empar Pons. 2005. "The Effects of Personality on Earnings." *Journal of Economic Psychology* 26 (3):363–384.
- Oldham, John M, Andrew E Skodol, and Donna S Bender. 2009. *Essentials of Personality Disorders*. American Psychiatric Pub.
- Papageorge, Nicholas W, Victor Ronda, and Yu Zheng. 2016. "The Economic Value of *Breaking Bad*: Schooling, Misbehavior and the Labor Market." Mimeo, The Johns Hopkins University.
- Paulson, Anna L, Robert M Townsend, and Alexander Karaivanov. 2006. "Distinguishing Limited Liability from Moral Hazard in a Model of Entrepreneurship." *Journal of Political Economy* 114 (1):100–144.
- Prada, María F. and Sergio Urzúa. 2016. "One Size Does Not Fit All: Multiple Dimensions of Ability, College Attendance, and Earnings." Forthcoming, *Journal of Labor Economics*.
- Rauch, Andreas and Michael Frese. 2007. "Let's Put the Person Back into Entrepreneurship Research: A Meta-Analysis on the Relationship between Business Owners' Personality Traits, Business Creation, and Success." *European Journal of Work and Organizational Psychology* 16 (4):353–385.
- Rees, Hedley and Anup Shah. 1986. "An Empirical Analysis of Self-Employment in the UK." *Journal of Applied Econometrics* 1 (1):95–108.
- Roberts, Brent, Joshua J Jackson, Angela L Duckworth, and Katherine Von Culin. 2011. "Personality Measurement and Assessment in Large Panel Surveys." *Forum for Health Economics and Policy* 14 (3):9.

- Roy, Andrew Donald. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (2):135–146.
- Rustichini, Aldo, Colin G DeYoung, Jon E Anderson, and Stephen V Burks. 2016. "Toward the Integration of Personality Theory and Decision Theory in Explaining Economic Behavior: An Experimental Investigation." *Journal of Behavioral and Experimental Economics* 64:122–137.
- Savelyev, Peter A. 2010. "Conscientiousness, Education, and Longevity of High-Ability Individuals." Mimeo, University of Chicago.
- Schumpeter, Joseph A. 1949. *Change and the Entrepreneur: Postulates and Patterns for Entrepreneurial History*. Harvard University Press.
- Shane, S, N Nicolaou, L Cherkas, and T Spector. 2010. "Openness to Experience and Opportunity Recognition: Evidence of a Common Genetic Etiology." *Human Resource Management* 29:291–303.
- Shane, Scott and Nicos Nicolaou. 2013. "The Genetics of Entrepreneurial Performance." *International Small Business Journal* 31 (5):473–495.
- Störmer, Susi and René Fahr. 2013. "Individual Determinants of Work Attendance: Evidence on the Role of Personality." *Applied Economics* 45 (19):2863–2875.
- Urzua, S. 2008. "Racial Labor Market Gaps The Role of Abilities and Schooling Choices." *Journal of Human Resources* 43 (4):919–971.
- Van der Sluis, Justin, Mirjam Van Praag, and Wim Vijverberg. 2008. "Education and Entrepreneurship Selection and Performance: A Review of the Empirical Literature." *Journal of Economic Surveys* 22 (5):795–841.
- Vandenberghe, Christian, Sylvie St-Onge et al. 2008. "An Analysis of the Relation Between Personality and the Attractiveness of Total Rewards Components." *Relations Industrielles/Industrial Relations* 63 (3):425–453.
- Viinikainen, Jutta, Katja Kokko, Lea Pulkkinen, and Jaakko Pehkonen. 2010. "Personality and Labour Market Income: Evidence from Longitudinal Data." *Labour* 24 (2):201–220.
- Wichert, Laura and Winfried Pohlmeier. 2009. "Female Labor Force Participation and the Big Five." Mimeo, University of Konstanz Department of Economics.
- Willis, Robert J. and Sherwin Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy* 87 (5):S7–S36.
- Zhang, Zhen, Michael J Zyphur, Jayanth Narayanan, Richard D Arvey, Sankalp Chaturvedi, Bruce J Avolio, Paul Lichtenstein, and Gerry Larsson. 2009. "The Genetic Basis of Entrepreneurship: Effects of Gender and Personality." *Organizational Behavior and Human Decision Processes* 110 (2):93–107.
- Zhao, Hao and Scott E Seibert. 2006. "The Big Five Personality Dimensions and Entrepreneurial Status: A Meta-Analytical Review." *Journal of Applied Psychology* 91 (2):259.
- Zhao, Hao, Scott E Seibert, and G Thomas Lumpkin. 2010. "The Relationship of Personality to Entrepreneurial Intentions and Performance: A Meta-Analytic Review." *Journal of Management* 36 (2):381–404.

9 Tables and Figures

Table 1: The Big 5 Personality Traits

Personality Trait	Associated With Being:
Openness to New Experiences	◇ Creative, imaginative, intelligent, curious, broad-minded, sophisticated and adventurous.
Conscientiousness	◇ Organized, responsible, hardworking and <i>not</i> careless.
Extraversion	◇ Outgoing, friendly, lively, active and talkative.
Agreeableness	◇ Helpful, warm, caring, softhearted and sympathetic.
Neuroticism	◇ Moody, worrying, nervous and <i>not</i> calm.

Description of the big 5 personality traits.

Table 2: SUMMARY STATISTICS

	Analysis Sample	Paid Employment	Self Employment	Δ (Self–Paid)	<i>p</i> -values
Earnings (2004)	\$78,328.06	\$74,687.57	\$93,988.64	\$19,301.07	0.008
Median Earnings (2004)	\$57,500.00	\$58,500	\$58,402	\$1,000.00	0.757
Assets in 1995	\$ 119,748.30	\$100,274.40	\$203,520.70	\$103,246.30	0.000
High school degree	0.19	0.19	0.20	0.01	0.674
Some college	0.27	0.27	0.26	-0.01	0.835
College graduate	0.25	0.25	0.25	0.00	0.931
Age	50.09	49.35	53.25	3.90	0.000
Married	0.79	0.78	0.83	0.05	0.136
No. of children	2.19	2.19	2.19	0.00	0.989
Spouse educ. (years)	14.46	14.53	14.17	-0.36	0.082
Spouse employed (1995)	0.66	0.64	0.72	0.07	0.110
Fluid Cognitive Ability	0.35	0.36	0.32	-0.04	0.618
Openness (1995)	3.07	3.06	3.13	0.07	0.083
Openness (2004)	2.97	2.95	3.06	0.12	0.008
Conscientiousness (1995)	3.40	3.40	3.44	0.04	0.232
Conscientiousness (2004)	3.46	3.46	3.48	0.02	0.580
Extraversion (1995)	3.14	3.12	3.25	0.14	0.004
Extraversion (2004)	3.04	3.02	3.16	0.14	0.005
Agreeableness (1995)	3.29	3.28	3.31	0.02	0.602
Agreeableness (2004)	3.24	3.23	3.30	0.07	0.102
Neuroticism (1995)	2.16	2.18	2.10	-0.08	0.152
Neuroticism (2004)	2.02	2.03	1.99	-0.04	0.453

Summary statistics for the analytic sample of 896 individuals, of whom 167 (19%) are self-employed.

Table 3: SECTOR EARNINGS AND SECTOR CHOICE

	Earnings		Sector Choice
	SE	PE	
Openness (2004)	-0.27 (0.34)	0.06 (0.05)	0.25 (0.12)
Conscientiousness (2004)	-0.1 (0.3)	0.13 (0.06)	-0.07 (0.12)
Extraversion (2004)	0.3 (0.31)	0.06 (0.05)	0.13 (0.11)
Agreeableness (2004)	-0.38 (0.29)	-0.18 (0.05)	-0.03 (0.11)
Neuroticism (2004)	0.02 (0.21)	0.05 (0.04)	0.03 (0.09)
Cognition	-0.14 (0.14)	0.12 (0.03)	0.04 (0.06)
Years of education	0.18 (0.06)	0.07 (0.01)	-0.004 (0.03)
Age	-0.02 (0.02)	0.008 (0.003)	0.04 (0.007)
Married	0.18 (0.34)	0.21 (0.06)	0.58 (0.27)
Number of kids	.	.	-0.06 (0.04)
Spouse education	.	.	-0.03 (0.02)
Spouse employed (1995)	.	.	0.22 (0.11)
Observations	169	727	896

OLS regressions of log earnings by sector (where [SE] refers to self-employment and [PE] refers to paid employment) along with probit estimates for sector choice, where the outcome variable is an indicator for 2004 self-employment. Standard errors are in parentheses.

Table 4: STRUCTURAL PARAMETER ESTIMATES

Panel A:	Earnings		Self Utility
Relative Returns	Self Emp.	Paid Emp.	Parameters
Openness	-0.24 (0.10)	0.07 (0.05)	0.79 (0.22)
Conscientiousness	-0.13 (0.11)	0.10 (0.05)	-0.10 (0.19)
Extraversion	0.26 (0.07)	0.05 (0.04)	-0.20 (0.18)
Agreeableness	-0.32 (0.09)	-0.16 (0.03)	-0.12 (0.20)
Neuroticism	0.02 (0.04)	0.09 (0.04)	0.23 (0.24)
Years of education	0.13 (0.01)	0.06 (0.01)	-0.07 (0.04)
Age	-0.02 (0.01)	0.01 (0.003)	0.06 (0.01)
Married	0.18 (0.16)	0.18 (0.06)	0.51 (0.40)
Cognition	-0.19 (0.06)	0.08 (0.03)	0.41 (0.18)
Number of kids	.	.	-0.10 (0.10)
Spouse education	.	.	-0.04 (0.02)
Spouse employed (1995)	.	.	0.19 (0.24)
Utility Weight	.	.	4.84e-05 (9.13e-06)
Constant	.	8.68 (0.09)	-2.72 (0.70)
Panel B:			
Business Ideas, Technology, Credit Constraints and Assets			
Business Idea Constant 1	6.67 (1.29)	Technology	0.14 (0.04)
Business Idea Constant 2	10.17 (0.20)	Credit Constraints	1.94 (1.67)
Business Idea Variance 1	0.94 (0.62)	1995 Assets	1.01 (0.24)
Business Idea Variance 2	0.19 (0.12)	Assets Variance	4.60 (1.39)
Mixture Probability	0.79 (0.01)		
Panel C:			
Earnings Shocks	Self Emp. (e^y)	Paid Emp. (e^w)	
e^{θ}	0.05 (0.05)	0.06 (0.02)	
\tilde{A}	0.04 (0.03)	0.03 (0.01)	
e^a	0.07 (0.04)	0.05 (0.02)	
Variance	1.25 (0.11)	0.58 (0.02)	

Parameter estimates along with standard errors in parentheses.

Table 5: PREFERENCES VERSUS EARNINGS

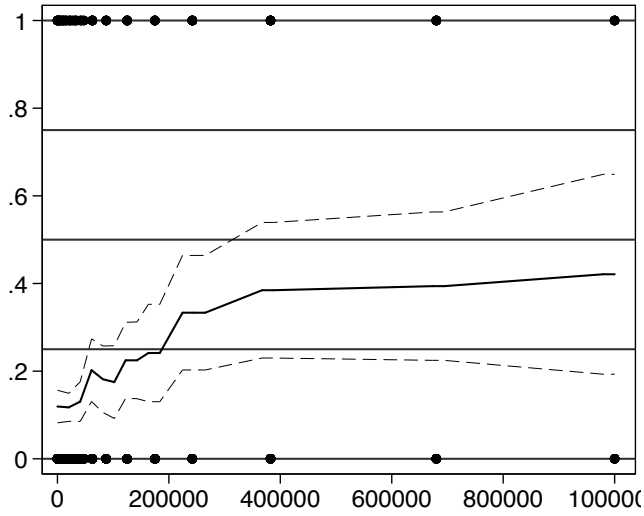
Characteristic	Highest Relative Payoff in SELF or PAID?		Misalignment?
	Earnings	Utility	
Education	SELF	PAID	✓
Age	PAID	SELF	✓
Married	SAME	SELF	
Cognition	PAID	SELF	✓
Openness	PAID	SELF	✓
Conscientiousness	PAID	PAID	
Extraversion	SELF	PAID	✓
Agreeableness	PAID	PAID	
Neuroticism	PAID	SELF	✓

For individual characteristics, we identify whether the marginal payoff (earnings and utility) is higher in paid or self-employment and then add a checkmark if there is a misalignment. For example, openness to new experiences leads to higher relative earnings in paid-employment, but higher utility in self-employment, leading to a misalignment.

Table 6: POLICY SIMULATIONS AND SELF EMPLOYMENT ENTRY AND EARNINGS

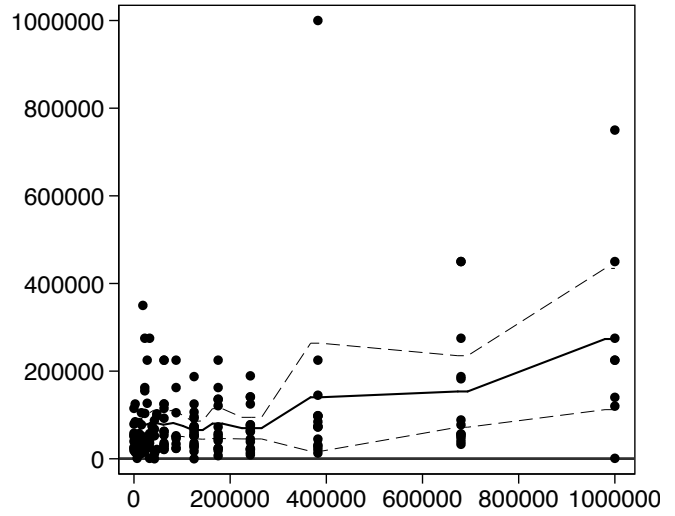
Counterfactual Policy Simulation	Entry Probability	% Change: Average SE Earnings	% Change: Average Val. of Bus. Ideas
Baseline	0.19	0.00	0.00
No Credit Constraints	0.22	0.02	<0.00
Preferences Do Not Affect Entry	0.13	0.30	0.19
Both	0.15	0.26	0.18
Subsidy	0.37	-0.46	-0.41
No Preferences and Subsidy	0.18	0.09	0.09

This table shows percent changes in entry probability, average self-employment earnings and the average value of realized business ideas (θ_i) under counterfactuals where: (i) there are no credit constraints, (ii) preferences do not affect entry (agents maximize earnings), (iii) there are no credit constraints and preferences do not affect entry, (iv) there is a blanket subsidy of \$25,000 for any individual who starts a business and (v) there is a blanket subsidy of \$25,000 for any individual who starts a business and preferences do not affect entry. In computing the percent change in self-employment earnings due to a counterfactual subsidy, we use earnings net of the subsidy.



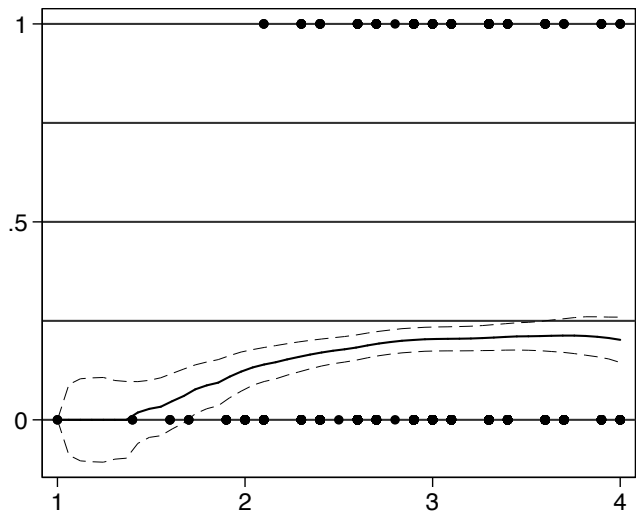
--- 95% CI
 — lpolynomial smooth: Self-employment
 • Self-employment

(a) 1995 Assets and Self Employment



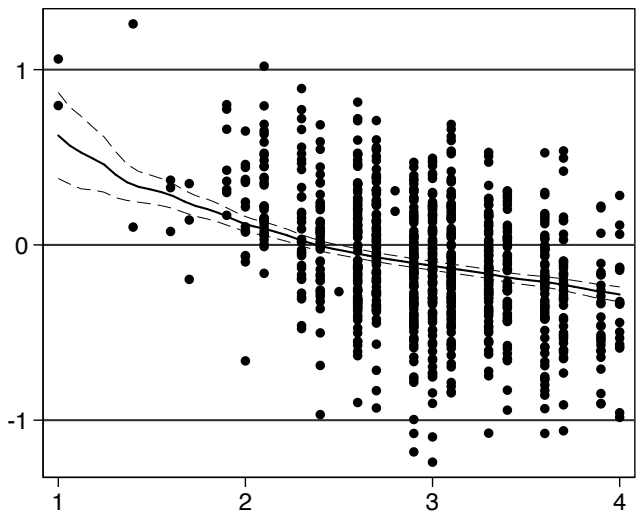
--- 95% CI
 — lpolynomial smooth: Earnings
 • Earnings

(b) 1995 Assets and Self Employment Earnings



--- 95% CI
 — lpolynomial smooth: Self-employment
 • Self-employment

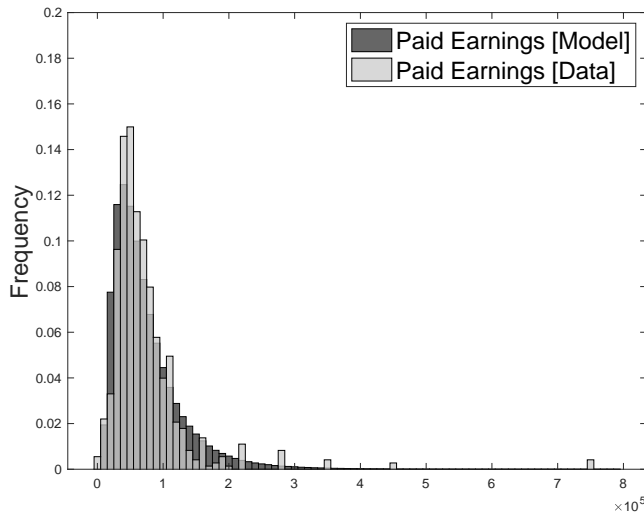
(c) Openness and Self Employment



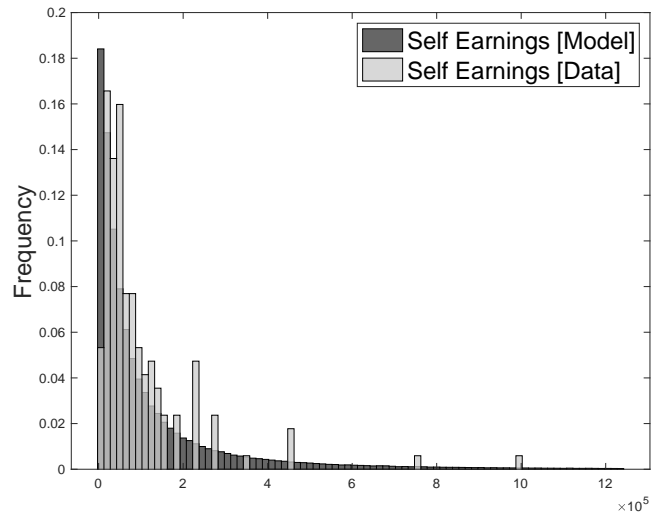
--- 95% CI
 — lpolynomial smooth: Relative log SE Earnings
 • Relative log SE Earnings

(d) Openness and Self Employment Earnings

Figure 1: EMPIRICAL PATTERNS. Panel 1(a) shows that self-employment rises with 1995 assets. Panel 1(b) shows that earnings rise with assets at high asset levels. Panel 1(c) shows a positive relationship between self-employment and openness; and Panel 1(d) depicts a negative relationship between openness and self-employment earnings.



(a) Paid Employment Model Fit



(b) Self Employment Model Fit

Figure 2: MODEL FIT. Earnings histograms from the data and predicted using estimated model parameters for individuals in paid employment and in self-employment.

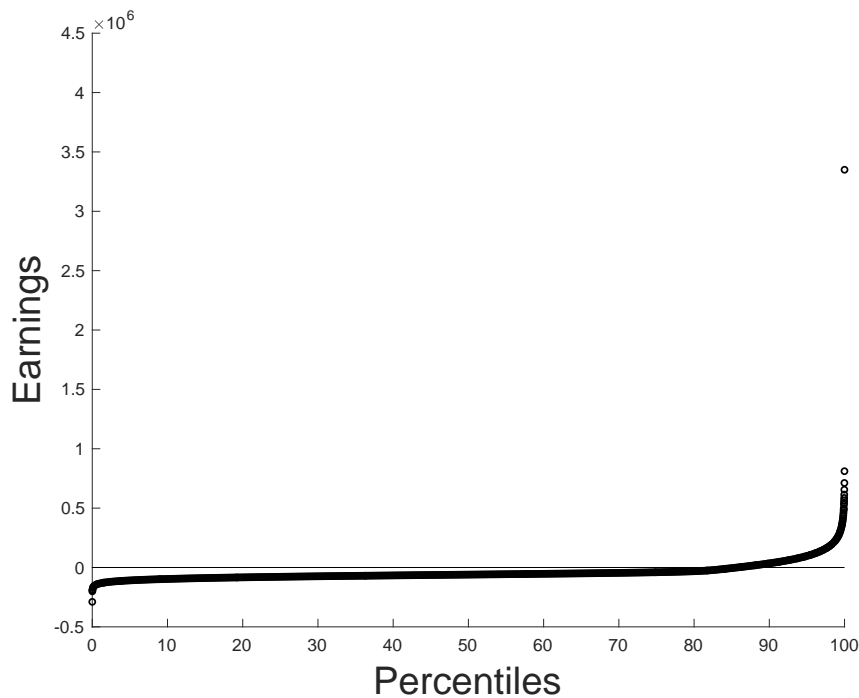
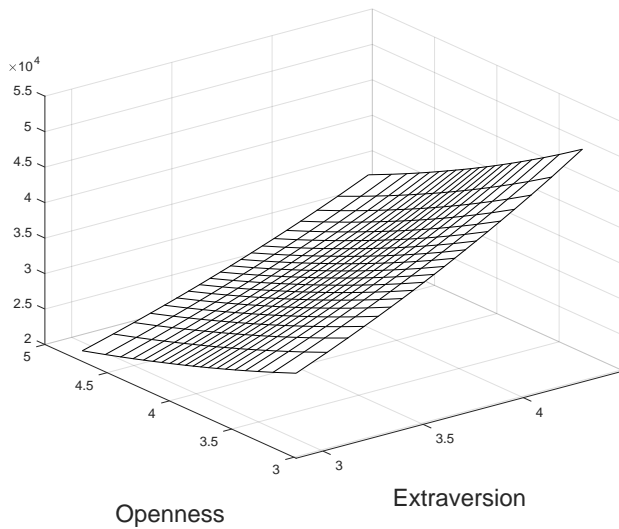
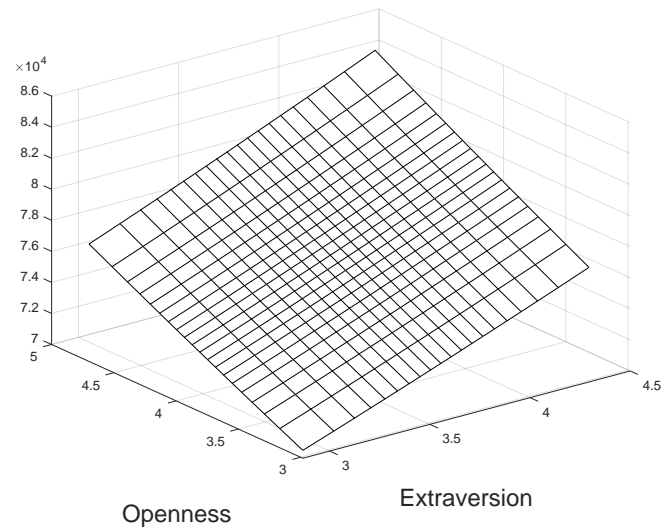


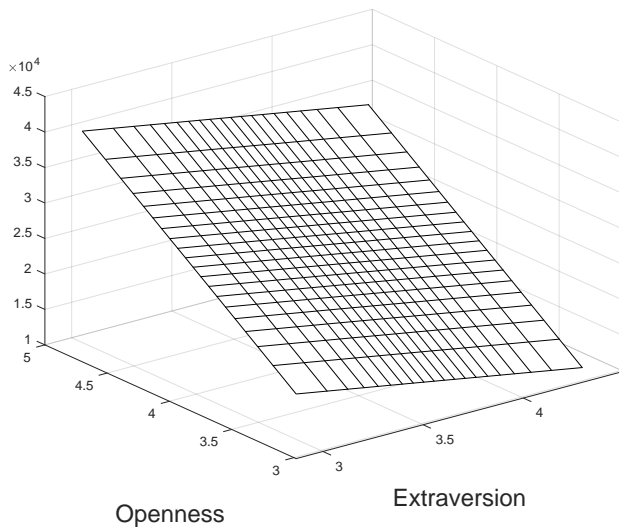
Figure 3: PERCENTILES OF SIMULATED EXPECTED EARNINGS DIFFERENTIALS. For each individual in the sample, expected self- and paid-employment earnings are drawn 2,500 times. All draws are ordered and plotted against their corresponding percentile. The x -axis is the tenth-percentile and the y -axis is earnings (in levels). The figure shows that most simulated workers (80%) would expect to lose money in self-employment. The simulation also illustrates why average earnings are high in self-employment: there is the possibility of an extremely high business idea draw.



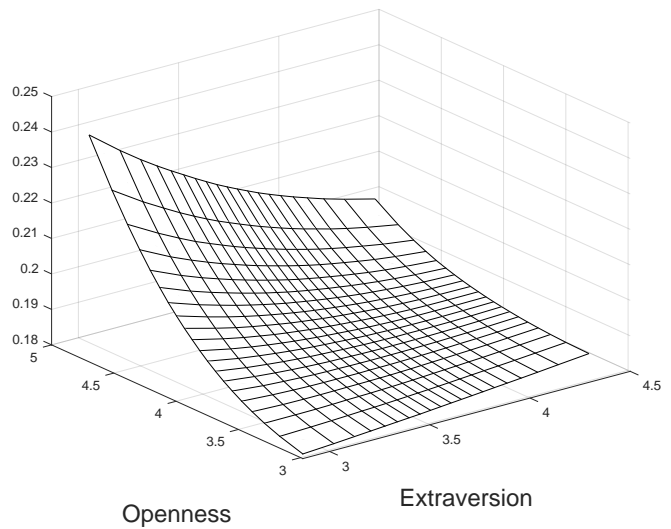
(a) Self Employment Earnings



(b) Paid Employment Earnings

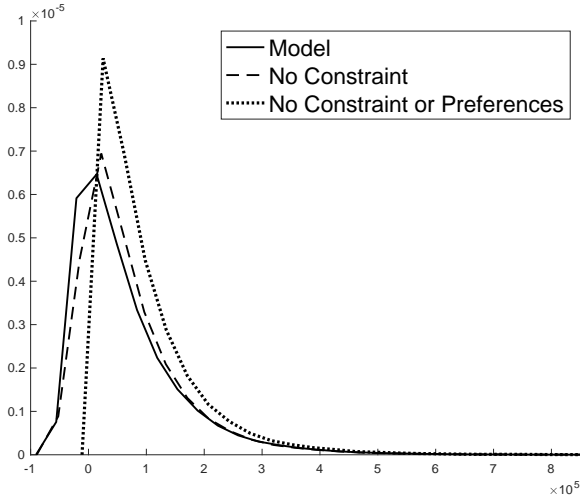


(c) Relative SE Utility

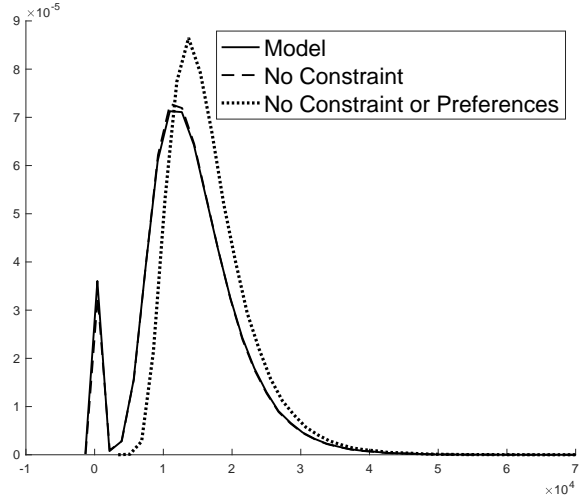


(d) Self Employment Entry

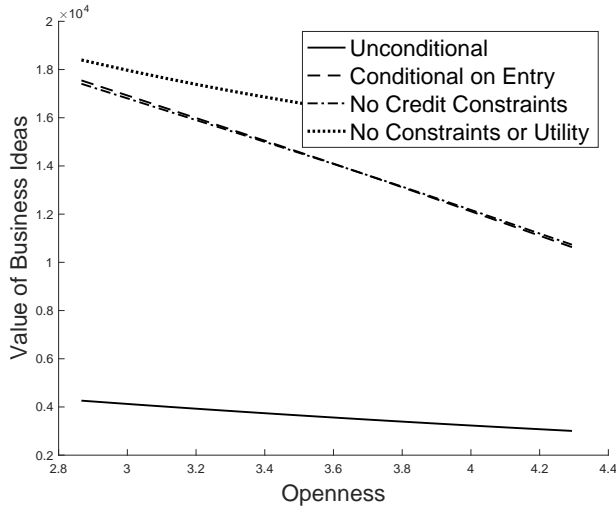
Figure 4: PERSONALITY AND SELF EMPLOYMENT: We simulate expected earnings for self- and paid employment, utility and entry probability where Openness to New Experiences and Extraversion are set equal to each combination of deciles for the sub-sample of individuals used in our analysis



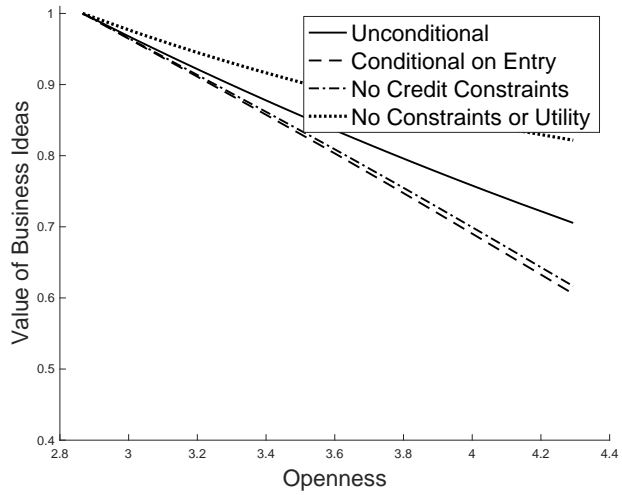
(a)



(b)

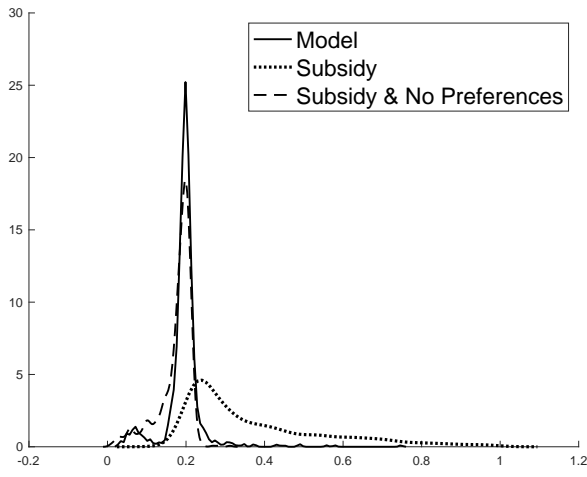


(c)

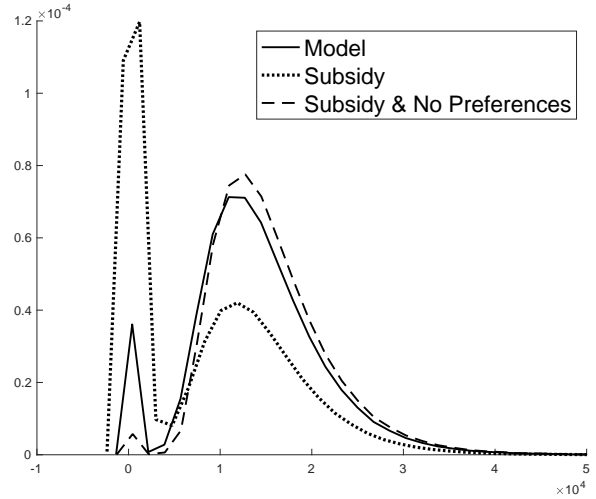


(d)

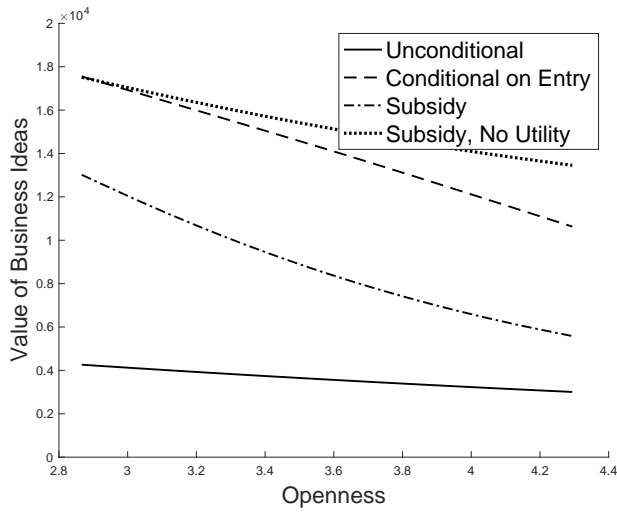
Figure 5: COUNTERFACTUAL SIMULATIONS — REMOVING IMPEDIMENTS. Panels 5(a) and 5(b) show simulated relative earnings and values of realized business ideas, respectively (i) using estimated model parameters (ii) absent a credit constraint and (iii) absent the credit constraint and where preferences play no role. Panel 5(c) plots the simulated relationship between openness and business ideas using model parameters (i) unconditionally, (ii) conditional on entry, (iii) absent credit constraints and (iv) absent credit constraints if preferences play no role. Panel 5(d) repeats the exercise of Panel 5(c) normalizing the value of business ideas to 1 at the lowest decile of Openness for ease of comparison.



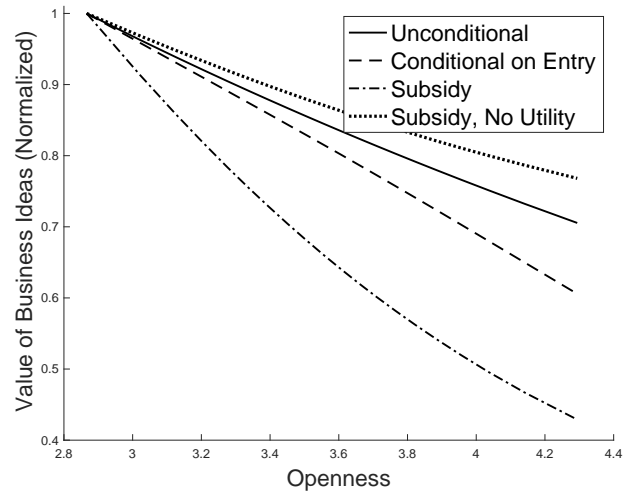
(a)



(b)

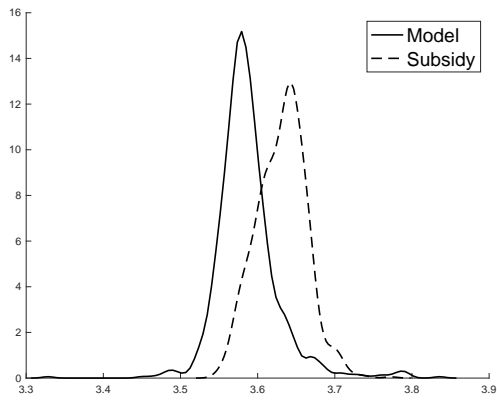


(c)

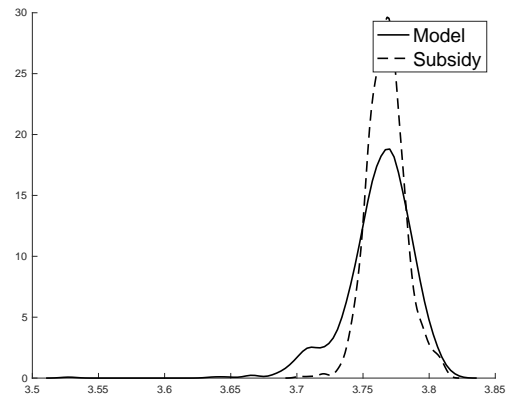


(d)

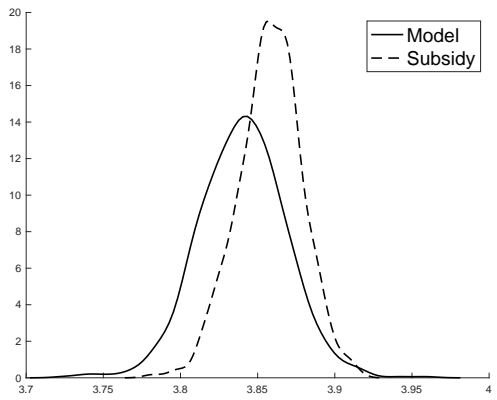
Figure 6: COUNTERFACTUAL SIMULATIONS — SUBSIDY. Panels 6(a) and 6(b) show simulated entry probability and values of realized business ideas, respectively (i) using estimated model parameters (ii) under a subsidy and (iii) under a subsidy and where preferences play no role Panel 6(c) plots the simulated relationship between openness and business ideas using model parameters unconditionally, conditional on entry, in the presence of a subsidy and with a subsidy if preferences play no role. Panel 6(d) repeats the exercise normalizing the value of business ideas to 1 at the lowest decile of Openness for ease of comparison.



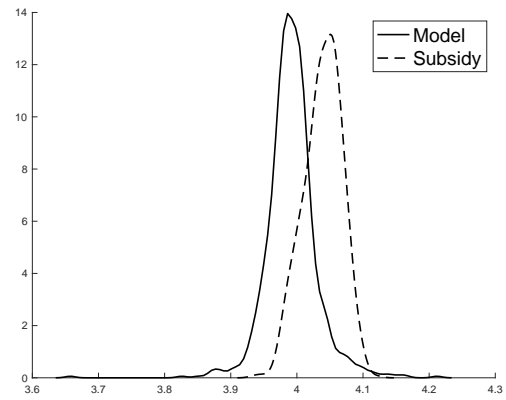
(a) Openness



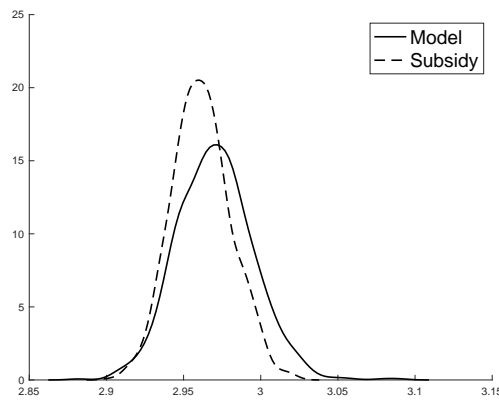
(b) Conscientiousness



(c) Extraversion



(d) Agreeableness



(e) Neuroticism

Figure 7: PERSONALITY AND SUBSIDIES. We plot the distribution of personality traits among individuals choosing self-employment simulated from the model and then under the counterfactual where entrepreneurship is subsidized (\$25,000 for all small businesses).

Appendix A Appendix

Appendix A.1 Identification

We are interested in identifying the distributions of five latent personality traits, as measured by the Big 5.⁵³ The aim is to show that we can use observed measurements of personality to identify the joint distribution of the latent factors along with the parameters mapping latent factors to observed outcomes.

To simplify exposition, suppose there are only two personality traits, each with two measurements (scores on personality assessments). Express the measurement of trait $j \in \{1, 2\}$ for agent i at time $t \in \{95, 04\}$ as follows:

$$\begin{aligned} C_{i1,95} &= f_{i1} + \epsilon_{i,95,1}^C \\ C_{i1,04} &= d_{1,04}^C f_{i1} + \epsilon_{i,04,1}^C \\ C_{i2,95} &= f_{i2} + \epsilon_{i,95,2}^C \\ C_{i2,04} &= d_{2,04}^C f_{i2} + \epsilon_{i,04,2}^C \end{aligned} \tag{22}$$

These expressions reflect two assumptions. One, each personality assessment is dedicated to a single underlying trait. Two, for each underlying trait, we normalize the 1995 factor loading (the parameter linking the underlying trait to the observed assessment) to 1.

For every individual in the sample, we also observe earnings (in either paid or self-employment.) To further simplify exposition, write log earnings as:

$$\ln(y_i) = \eta_1 f_{i1} + \eta_2 f_{i2} + \nu_i^y. \tag{23}$$

Finally, we assume that the ϵ are independent of each other and of ν_i^y . However, we do not assume that $f_{i1} \perp f_{i2}$ — i.e., we permit underlying latent traits to be correlated.

The aim is to show that we can identify the κ , which requires identification of the joint distribution of the underlying factors f_{i1} and f_{i2} . Under stated assumptions (independence of error terms, dedicated measurements and the normalizations), it holds that:

$$\begin{aligned} d_{1,04}^C &= \frac{\text{Cov}(C_{i1,04}, \ln(y_i))}{\text{Cov}(C_{i1,95}, \ln(y_i))} \\ d_{2,04}^C &= \frac{\text{Cov}(C_{i2,04}, \ln(y_i))}{\text{Cov}(C_{i2,95}, \ln(y_i))} \end{aligned} \tag{24}$$

⁵³We adapt identification arguments presented, for example, in Urzua (2008).

This allows us to obtain the variance and covariance of each factor:

$$\begin{aligned}
\text{Var}(f_{i1}) &= \frac{\text{Cov}(C_{i1,95}, C_{i1,04})}{d_{1,04}^C} \\
\text{Var}(f_{i2}) &= \frac{\text{Cov}(C_{i2,95}, C_{i1,04})}{d_{2,04}^C} \\
\text{Cov}(f_{i1}, f_{i2}) &= \text{Cov}(C_{i1,95}, C_{i2,95})
\end{aligned} \tag{25}$$

Finally, η_1 and η_2 are identified from the covariance of $\ln(y_i)$ and $C_{i1,95}$ and $C_{i2,95}$, respectively.

Identification of the remaining unobservables in the model is less straightforward. Self-employment and paid-employment earning shocks are dependent on both the business idea and asset growth disturbances (e^θ and e^a). As a result, it follows that both disturbances also influence the individual’s sector choice. Identification of the model relies on the identification of the asset disturbance. To see that, note that the model without the asset growth disturbance (e^a) is a modified version of the Roy model (Roy, 1951) and standard identification arguments would apply (see e.g., Heckman and Honore (1990)). Identification of the asset equation is possible since the asset disturbance enters non-linearly in the credit constraint equation (see especially equations (16) and (19)). Otherwise it would not be possible to distinguish between the two disturbance shocks.

To assess robustness, we have tried some alternatives, which includes estimating the asset growth equation outside of the model using average asset growth from 1995 to 2004, though this may induce endogeneity through reverse causality, which is what we aim to avoid. We can also use 1995 assets as a noisy measure of 2004 assets. These alternatives do not change main qualitative results appreciably. One important exception is that using 1995 assets leads to a noisier estimate of the credit constraint parameter λ . This likely occurs because 1995 assets are too low relative to predicted 2003 assets. This leads to an under-estimation of the importance of credit constraints in reducing entry. Finally, it is worth mentioning that, similar to predicted 2003 assets, observed 2004 assets exhibit far higher variance than 1995 assets, which is noteworthy since we do not use 2004 assets as a moment to be matched, i.e., the rise in asset variance over time is matched “out-of-sample.”

Appendix A.2 Estimation Algorithm and Likelihood

The simulation procedure begins as follows: we draw a block matrix (denoted B) of size $K \times I \times (J + 2)$ from a standard normal distribution. Recall that J is the number of personality traits and K is the number of draws per individual. We need a block matrix of size $J + 2$ since we draw not only J personality traits, but also unobservables for the mixed-normal distribution of business ideas. We draw B once. Next, at each parameter

suggestion $\Phi^{(g)}$ and for each individual i , we compute expected earnings in paid employment (denoted $w_{ik}^{(g)}$), expected earnings in self-employment (denoted $y_{ik}^{(g)}$) and the resulting sector choice (denoted $d_{ik}^{(g)}$). For earnings and choices, the superscript (g) indexes the parameter suggestion and the subscript ik refers to the k -th draw of individual i .

The simulation of earnings and sectoral choice occurs in several steps. Using parameters $\Xi_f^{(g)}$, we simulate vectors of latent factors $f_{ikj}^{(g)}$, $j \in \{1, \dots, J\}$ for each individual i and draw k . Similarly, we use the parameters $\mu_{\theta,1}^{(g)}$, $\mu_{\theta,2}^{(g)}$, $\sigma_{\theta,1}^{2(g)}$, $\sigma_{\theta,2}^{2(g)}$ and p^θ to simulate a business draw for each individual i and draw k , which we denote $\theta_{ik}^{(g)}$. From here, we can determine whether or not each individual-draw pair is credit-constrained using equation (3) suitably modified to permit multiple draws. In particular, individual i with draw k and at parameters (g) is credit-constrained if:

$$\theta_{ik}^{(g)} \psi_{ik}^{(g)} > \frac{r}{\alpha^{(g)}} (\lambda^{(g)} A_i)^{1-\alpha^{(g)}}. \quad (26)$$

Note that the k subscript is omitted from α , which remains constant across all K draws. Moreover, assets A_i , which are data, and the interest rate r (set to 1.1 for this analysis) do not change with draws or with suggested parameters (g) . $\theta_{ik}^{(g)}$, however, is different for each individual i , draw k and parameter suggestion (g) .

Once it is clear which individuals are credit-constrained, we can compute $y_{ik}^{(g)}$ for each individual, using r , A_i , $\alpha^{(g)}$ and $\lambda^{(g)}$ when the credit constraints are binding and r , $\alpha^{(g)}$ and $\theta_{ik}^{(g)}$ when they are not binding. Similarly, we compute utility $u_i^{SE(g)}$ and paid earnings $w_i^{(g)}$ using parameter suggestions. Then, using equation (19), we compute a sector choice for each individual-draw pair, denoting this $d_{ik}^{(g)}$. In what follows, we use $f_{ikj}^{(g)}$, $w_{ik}^{(g)}$, $y_{ik}^{(g)}$ and $d_{ik}^{(g)}$ to construct the likelihood.

The likelihood function consists of several components. Given the assumption that earnings shocks are normally distributed, we form the earnings portion of the likelihood using the normal density function, which for individual i , draw k and parameter suggestion g we denote $h(y_{ik}^{(g)})$ and $h(w_{ik}^{(g)})$ for self-employment wage density and paid employment wage density respectively. Next, given assumptions on the normality of the measurement error in latent traits, we can also derive the density function for personality measurements for each individual i , draw k and parameter vector g , denoting this $h(M_{ik}^{(g)})$. Then, we must average these, though these averages are conditional on the relevant sector being chosen for a given draw:

$$L_i^{y(g)} \equiv \frac{1}{K_{i,SE}^{(g)}} \sum_{k=1}^{K_{i,SE}^{(g)}} \left[h\left(y_{ik}^{(g)}\right) \times h\left(M_{ik}^{(g)}\right) \mid d_{ik}^{(g)} = \text{SE} \right] \quad (27)$$

and

$$L_i^{w(g)} \equiv \frac{1}{K_{i,PE}^{(g)}} \sum_{k=1}^{K_{i,PE}^{(g)}} \left[h \left(w_{ik}^{(g)} \right) \times h \left(M_{ik}^{(g)} \mid d_{ik}^{(g)} = \text{PE} \right) \right]. \quad (28)$$

In the above equations, $K_{i,SE}^{(g)}$ denotes the number of draws for which individual i at parameter draw (g) chooses self-employment. Similarly, $K_{i,PE}^{(g)}$ denotes the number of draws for which individual i at parameter draw (g) chooses paid employment. $L_i^{y(g)}$ and $L_i^{w(g)}$ are the product of average earnings densities for each sector and average personality trait densities, conditional on a sector being chosen. Therefore, they are a weighted average of each individual's likelihood contribution, where the average is taken over the subset of the K draws where the individual chooses the relevant sector at draw k .

Next, we weight likelihood contributions by the probability that the model predicts that a sector is chosen by a given individual. We denote this probability \tilde{P}_i , defined as the number of times that the individual chooses self-employment given K draws:

$$\tilde{P}_i = \frac{K_{i,SE}^{(g)}}{K}. \quad (29)$$

Then, the likelihood contribution for individual i and draw k will be given by:

$$L_i^{(g)} = \left[\tilde{P}_i^{(g)} \times L_i^{y(g)} \right]^{d_{it}=\text{SE}} \left[\left(1 - \tilde{P}_i^{(g)} \right) L_i^{w(g)} \right]^{d_{it}=\text{PE}}, \quad (30)$$

where d_{it} is the observed sector choice so that, for each individual, the contribution to the likelihood is only a function of the probability the model predicts their observed sector is chosen, multiplied by the average of the product of the earnings density in that sector and personality traits density, where the average is conditional on the model predicting that sector.

After constructing $L_i^{(g)}$ for each individual i , we take the log of each individual's contribution and then sum over individuals to obtain the log-likelihood:

$$l^{(g)} = \sum_{i=1}^I \log \left(L_i^{(g)} \right). \quad (31)$$

We evaluate $l^{(g)}$ at different values in the parameter space, indexing these suggestions by (g) and, using both simplex and gradient methods, search until a maximum is found.