Entrepreneurial Human Capital and Entrepreneurial Dynamics:

Evidence from a Natural Experiment*

Russell Toth,†

February, 2012

Abstract

This paper tests a model in which entrepreneurship-specific human capital accumulation through learning-by-doing is the key factor behind entrepreneurial dynamics. I derive testable predictions of the theory that allow it to be distinguished from alternatives. I then exploit the 1997-98 Indonesian financial crisis as a natural experiment that provides exogenous variation in entry into self-employment amongst a relatively high-ability cohort of individuals. Consistent with the model of human capital accumulation, entrepreneurial activity is robustly persistent, and the dynamic improvement in returns exceeds what could be reasonably expected in the absence of skill accumulation.

Key words: Entrepreneurial human capital, learning-by-doing, occupational choice, subsistence and opportunity entrepreneurs

JEL Codes: O12, L26, J24, D92

*I am grateful to David Easley, Chris Barrett, Ted O'Donoghue and Viktor Tsyrennikov for their guidance and feedback. I also thank Jim Berry, AV Chari, Corey Lang, Josh Lerner, Maximilian Mihm, Ervin Starr, discussants Jenny Aker and L.G. Thomas, and seminar audiences at Australian National University, Cornell University, Harvard Business School, Harvard Business School International Research Conference (2011), The University of Sydney, University of Waterloo, and NEUDC (MIT-Sloan, 2010) for useful feedback. I thank Andre Syafroni, Maria Wihardja and KPPOD for assistance with fieldwork. I acknowledge support from the Cornell University Institute for the Social Sciences, the Mario Einaudi Center for International Studies, the Cornell University Graduate School, and NSF Expeditions grant num. 0832782. All errors are my own.

†School of Economics, Room 363 Merewether Building H04, The University of Sydney, NSW 2006 Australia. Web: https://sites.google.com/a/cornell.edu/russelltoth/. Email address: russell.toth@sydney.edu.au

1 Introduction

The question of what drives entrepreneurial dynamics, particularly entry and enterprise growth, is of long-standing interest to policymakers and academics in both the advanced and developing world. Such issues have taken on a new urgency in the developing world due to the pressures and opportunities of globalization and an increased recognition of vibrant private-sector enterprise activity as a source of economic growth and poverty reduction. In the academic literature there has been a recent, burgeoning interest in the factors behind entrepreneurial dynamics and firm performance. This has partly been motivated by the large interest, and subsequent disappointment, in microfinance as a stimulant of widespread growth-oriented entrepreneurship (Banerjee, Duflo, Glennerster, and Kinnan [2009]; Karlan and Zinman [2010]), and also a growing interest in cross-country differences in firm productivity (Bloom, Mahajan, McKenzie, and Roberts [2010]).

While much existing literature has focused particularly on firm-level financing access constraints as the key factor behind enterprise dynamics and outcomes, there has been growing awareness that this might not tell the full story. Attention has begun to shift to the human capital and managerial skill and knowledge of the individual entrepreneur and the firm (see, e.g., Bloom, Mahajan, McKenzie, and Roberts [2010], Bruhn, Karlan, and Schoar [2010]), and the distinction between the much larger cohort of subsistence entrepreneurs and the significantly smaller cohort of higher-potential, transformational entrepreneurs (Schoar [2010]). While existing theories largely take the entrepreneurship-specific human capital of the entrepreneur as a fixed parameter and focus on other factors behind enterprise dynamics such as learning and credit-savings interactions (e.g., Jovanovic [1982], Buera [2009]), in this paper I derive and test predictions of a simple theory of entrepreneurship-specific human capital accumulation.

To test the theory and provide additional evidence on the returns to entrepreneurial experience, I exploit a uniquely-suitable natural experiment, the 1997-98 East Asian financial crisis. Key to identifying the stock of entrepreneurial human capital is the ability to disentangle it from other (generally unobservable) factors generating selection into entrepreneurial occupations, such as the role of a market opportunity, a "good idea," or unobserved ability or information. The crisis provides a plausibly unanticipated shock that generates increased entry into self-employment, which is orthogonal to a number of potential confounding factors. I focus on Indonesia, where the crisis had its most drastic effects among all countries. During the crisis period the returns to informal self-employment relative to private sector wage employment shift sharply in favor of self-employment for a significant cohort of individuals, providing an exogenous source of selection into self-employment. In addition, this effect is much more pervasive at the upper end of the earnings distribution (Thomas, Beegle, and Frankenberg [2000]), allowing us to focus on a higher-ability cohort of potential entrepreneurs.

We can think of this as a convenient natural laboratory in which to study the factors behind entrepreneurial dynamics. A cohort with higher potential to be transformational entrepreneurs, who might not normally consider running a business, are suddenly forced to do so. They generally start enterprises in the informal non-farm sector, which was relatively unaffected by the financial crisis relative to the formal sector. Many have little prior self-employment experience. This focus on high-ability types is important in light of recent literature pointing to significant heterogeneity in the entrepreneurial potential of individuals (e.g., Barrett, Bezuneh, Clay, and Reardon [2005], de Mel, McKenzie, and Woodruff [2010], Porta and Shleifer [2008], Schoar [2010]). Such work suggests that the subset of individuals with the potential to run relatively larger, growth-oriented enterprises is a relatively small subset of the entrepreneurial talent distribution. Additionally, much of the existing developing-country literature has focused on the self-employment experiences of low-ability individuals, or draws inferences on higher-ability types from samples that are subject to non-random selection and recall biases. Finally, while recessionary and crisis events in more developed economies also have been shown to increase self-employment, such cases are less useful for the purposes of this paper because the selection effect tends to be concentrated on lower-ability individuals, and is distorted by social safety nets such as unemployment insurance, severance packages, firing restrictions, and the like.

My primary dataset is the Indonesia Family Life Survey (IFLS), a panel survey which follows a nationally-representative sample of around 10,000 households and spans the crisis years. A particularly striking and puzzling finding in the data is that self-employment activity is remarkably persistent amongst those who enter self-employment during the crisis. This is true even in the years after the effects of the crisis have dissipated and the cohort of crisis entrants might have been expected to return to waged employment. About 78% of individuals who enter self-employment during the crisis are still running a business as their primary occupation 9 years after the crisis, whereas the usual 9-year persistence rate is around 46%.

I conduct more detailed empirical analysis in order to show that this persistence is robust. To do so I estimate counterfactual models of self-employment participation, to compare predicted self-employment participation to realized outcomes. I show that even after controlling for other covariates, in particular individuals' expected wage (as a measure of opportunity cost), individuals are estimated to be three to ten times more likely to engage in self-employment than under the counterfactual. I interpret this as evidence that after the initial, unexpected entry into self-employment, the individuals' stock of entrepreneurship-specific human capital has increased to the point of altering long-run occupational choice incentives. This explanation is also consistent with the jump in persistence (from 46% to 78%) amongst crisis-period entrants. If we make the natural assumption that the learning curve is concave and hence relatively steep early in one's entrepreneurial career, then we would expect

the largest jump in persistence amongst a cohort of individuals with little prior experience who are exogenously pushed into self-employment.

Since alternative theories also predict persistence, however, I also look at unique predictions of the theory of entrepreneurial human capital accumulation regarding earnings. Competing theories that take the stock of entrepreneurial skill as fixed generate sub-optimal initial physical capital allocations due to uncertainty over own-ability (Jovanovic [1982]) or credit-market constraints (Buera [2009]). Under such theories firm growth (and persistence) is input-driven, as the stock of labor and capital in the firm grow to match the ability endowment of the entrepreneur. However, such theories require certain patterns in the co-movement of inputs and returns to hold (under a reasonable specification of the production function). I show that the trajectory of returns is best explained by the human capital-acquisition theory, as it exceeds potential benchmarks including the dynamic expansion of inputs.

Finally, having provided evidence in support of the theory of entrepreneurial human capital accumulation, I proceed to directly estimate the causal effect of entrepreneurial experience on earnings. I use self-employment experience as a proxy for entrepreneurship-specific human capital acquisition through learning-by-doing, since underlying entrepreneurial ability is not directly observed. While a substantial literature in labor economics that similarly proxies human capital accumulation through years of exposure generally points to a positive effect of human capital on entry and earnings,[3] rigorous empirical evidence focused on human capital specific to entrepreneurship is much less common. In this case the separation between formal-sector labor market churning and outcomes in the informal sector is used as the key exclusion restriction in an instrumental variables setup. The evidence is supportive of the quantitative importance of the role of human capital acquisition, as experience is shown to strongly increase earnings, with the best estimate suggesting on average a 3% boost in net profit for each additional year of experience.

The paper makes a number of contributions to the literature. It provides new evidence and an explanation for the surprising persistence in entrepreneurship that we see amongst higher-ability individuals after the Indonesian financial crisis. It suggests that a theory of transformational entrepreneurship should reflect the role of learning-by-doing in driving entrepreneurial dynamics. It also provides evidence on the value of entrepreneurial learning-by-doing for the unique cohort of crisis-period entrants, providing some of the first estimates on the value of such human capital in the literature.

The paper proceeds as follows. I begin with a qualitative description of the setting, with a particular

[3]See Card [1999] for a review of studies taking years of education as a proxy for generalized human capital accumulation, and Angrist [1990] and Behrman and Rosenzweig [1999] for studies of human capital accumulation through work experience.

focus on the informal sector in Indonesia and the effects of the crisis, along with a discussion of entrepreneurial human capital, in Section 2. I then outline a simple, dynamic model of entrepreneurial selection, savings and consumption in Section 3, and derive testable predictions that can be used to distinguish it from competing theories. I provide descriptive evidence on the dataset and outline the identification strategies in Section 4. The empirical analysis has two foci: evidence in favor the theory of entrepreneurial human capital accumulation, and causal evidence on the effects of entrepreneurial experience on earnings. The results are presented in Section 6, while various threats to analysis and robustness checks are presented in Section 7. Section 8 concludes, while additional content appears in Appendix A.

2 The Setting, and Entrepreneurial Human Capital

2.1 Indonesia Background

Indonesia is the world's fourth-largest country by population, and the largest Muslim democracy, though civil society is relatively secularized. In 1970 it was one of the world's poorest countries by any measure. However, it enjoyed average economic growth of 4.5% per year between the mid-sixties until the 1997-98 Asian Financial Crisis, and was on the verge of joining the middle income countries. In 1998 GDP dropped by 14% at the height of the crisis. After the end of President Suharto's reign during the crisis, the country began a political transition, which has involved full, democratic elections, regulatory reform, and decentralization of power. By 2000 GDP growth recovered to 5% and was fairly steady around 5-6% until 2008.

Geographically, the country is spread out over thousands of islands in the world's largest archipelago. The country is highly diverse ethnically, religiously, linguistically and economically, yet unified by a common major language and national institutional structure. The island of Java, which contains the capital city of Jakarta, is the central economic hub. Per capita gross domestic product in purchasing power parity terms was $4000 USD in 2009, putting Indonesia at 155 in the world in this category (for comparison, the equivalent value for the United States is $46,400), while the value of the Gini coefficient was 39.4 in 2005, which is about average in international comparison (CIA [2012]).

2.1.1 Indonesian Labor Market Status and Trends

In most less-developed economies more than half of the workforce is engaged in operating or working in microenterprises,[4] which generate roughly half of GDP. A negligible proportion of such enterprises manage to grow beyond subsistence scale. In more developed economies the contribution of microenterprises to employment and GDP is closer to 15%, while an active and large small and medium enterprise sector that is absent in most developing countries contributes close to half of GDP. In Indonesia more than half of the workforce has typically been involved in working in or running micro, small or medium enterprises. The vast majority of such enterprises are informal sector firms with less than 10 employees.

The formal sector was expanding in Indonesia in the years leading up to the crisis. From the mid-1980s until the late 1990s agricultural employment declined from 55 to 41% of the workforce, while manufacturing employment increased from 8 to 19% over the same period. Workforce participation rates of women were also increasing in the years leading up the crisis, from about 30 to 37% in the urban sector, though there was a steadier level of participation in rural areas of around 55%. This increase in urban employment was enjoyed in both the wage and self-employment sectors. We also notice that women are much more likely to work as workers in household enterprises in rural areas, at a 20-30% rate. Overall we see that labor force participation is relatively stable leading up the crisis, with a small uptrend toward formal sector activity.

2.1.2 The 1997-98 Crisis

The study of the labor market and self-employment effects of the crisis is facilitated by the availability of two excellent micro-datasets, which is unusual for a developing country. SAKERNAS is a labor-force survey that is collected by the Indonesian government statistical service, BPS, and is a large-scale, cross-sectional labor force survey. The Indonesian Family Life Survey (IFLS), is a panel dataset that was collected over multiple years, including the years spanning the crisis. For the study of the effects of a large-scale economic crisis, the dataset is particularly exciting because it included rounds just before the crisis hit, in 1997, and then a one-year-later follow-up. The IFLS is the primary dataset that will be used in the subsequent analysis in this paper.

Smith, Thomas, Frankenberg, Beegle, , and Teruel [2002] and Thomas, Beegle, and Frankenberg [2000] provide evidence on the labor market effects of the crisis. It is broadly recognized that Indonesia was the country worst-hit by the crisis, and that it was an unexpected event. The primary direct victim of the crisis was the banking and financial-services sector, much of which was left out

[4]Indonesia's official enterprise size cohorts are defined as follows: microenterprise (1-4 workers), small enterprise (5-19 workers), medium enterprise (20-99 workers) large enterprise (100+ workers).

or reorganized. The banking sector fell into disarray, and this led to a shortening of credit. While some of the early post-crisis research suggested that the crisis caused massive unemployment, in fact this claim does not hold up in the micro-data. What we see instead is significant churning in occupational allocations, with one important movement being from private sector waged employment to self-employment. The government sector seems to be relatively well-sheltered from the effects of the crisis.

Consumer prices began to spiral upward in 1998, at the rate of 80% in that year according to CPI. Hence a number of price subsidies were removed, such as on rice, oil and some fuels. All of this uncertainty and economic pain led to the fall of President Suharto in May 1998, with multi-party elections and the return to relative stability in 1999. The shock to relative prices that the crisis brought about did have some beneficiaries – exporters, export producers and the like. Those producing services and non-tradeables likely did less well, though on the other hand the informal sector was also better-sheltered from the crisis, by being more independent from formal sector financial institutions ex ante.

The labor market and consumer effects were a derivative of the impacts on firms and the price rises. On average real wages collapsed by 40% between August 1997 and August 1998, and these effects reached most sectors of the economy. However, informal sector effects were less pervasive, particularly amongst rural, self-employed males. Of greater interest to the current study is the resulting *relative* price changes, as reported in Thomas, Beegle, and Frankenberg [2000]. In particular, there is strong evidence that the main relative price shock during the crisis was in expanding the relative margin between waged employment and self-employment. Self-employment broadly became relatively more attractive in comparison to private, waged employment, on the order of a 25-60% shift in relative returns depending on sector, gender and urban or rural location. In addition, this effect seems to be more strongly concentrated at the upper-end of the wage distribution (Smith, Thomas, Frankenberg, Beegle, , and Teruel [2002]; Thomas, Beegle, and Frankenberg [2000]), and we see the most significant occupational churning from private wage to self-employment at the upper end of the wage distribution.[5] This suggests that it might be relatively high-ability individuals who were induced to enter self-employment during the crisis.

Hence it appears that the crisis can be interpreted as a large, unexpected shock to the choice margin between private wage employment and self-employment, which hits the most able formal-sector workers the hardest. This was due both to a significant hit to private wage returns, along with the observation that informal-sector, self-employment activity was generally more sheltered from the crisis. This exogenous and unanticipated shift in the choice margin appears to have induced sectoral restructuring toward self-employment activity.

[5] Poppele, Sumarto, and Pritchett [1999] argue that the main effect of the crisis was on the urban elites.

2.2 Entrepreneurial Human Capital

Entrepreneurial human capital (EHC) constitutes specialized, high-level entrepreneurship-specific skills and knowledge, such as in selling, negotiating, product development, risk judgment (Shane [2003]) and entrepreneurial social capital. Above and beyond heterogeneous *ex ante* endowments of innate EHC, perhaps due to genetic inheritance or early upbringing (i.e., dynastic transitions), I hypothesize that EHC is significantly accumulated through *direct* exposure to entrepreneurial activity. Such entrepreneurial capabilities are distinct from other, generalized forms of human capital such as education, life experience, and experience in waged employment. A primary channel for acquiring EHC is learning-by-doing (i.e., running an enterprise, the focus of this paper).[6] EHC cannot be transacted in the marketplace separately from the individual endowed with it and public institutions for the transmission of EHC (such as the formal education system for wage work) are generally absent. Hence dynamic occupational selection incentives can play a crucial role in individuals' ability to accumulate EHC.

3 A Simple, Dynamic Model of Entrepreneurial Selection, Savings, and Consumption

I outline a simple, forward-looking model of individual occupational choice that captures the theoretical mechanism that I will test for in the data. The key feature of the theory is that it allows for entrepreneurial human capital accumulation through direct learning-by-doing. This is meant to capture the accumulation of *entrepreneurship-specific* human capital and business capital, such as in product development, marketing, risk judgment and business-relevant social network connections, through first-hand exposure to entrepreneurial activity. Much of the existing literature takes entrepreneurial human capital as fixed,[7] and then studies dynamics emanating from the physical capital accumulation (savings) choices of the entrepreneur,[8] or learning about the value of the unknown

[6] Other channels for EHC transmission that one might consider include transmission of EHC in the family (e.g., learning from one's parents, if they are entrepreneurs), or learning through work experience in another firm.

[7] The seminal, early reference on job- and occupation-specific human capital is Becker [1964]. Surprisingly little work has been done to formally extend such ideas to entrepreneurship, though less formal work exists in the economics literature in the work of Schultz [1980]; see also Klein and Cook [2006]. Two exceptions, though less general in scope, are Otani [1996] and Iyigun and Owen [1998].

[8] The literature focusing on occupational choice and the dynamic savings problem has primarily been motivated by an attempt to rationalize otherwise suprisingly strong inequalities in the aggregate wealth distribution. In such models individuals with (unobserved) high entrepreneurial skill have an incentive to save much more than others, which can generate significant wealth inequalities in a dynamic setup. See, e.g., Cagetti and Nardi [2006] and Buera [2009].

entrepreneurial ability endowment in a Bayesian learning framework.[9] In the empirical analysis I will argue that the learning-by-doing framework best matches the patterns in the data. Hence I begin by providing a simple formalization of the learning-by-doing framework, then derive testable predictions of the modeling frameworks that allow me to disentangle them in the data.

The agent is initially endowed with a stock of liquid wealth, W^0. While the agent can be thought of as capturing a household unit, the exposition will describe the model for a single individual. In each of two periods, $t = 0, 1$, the individual makes a choice between one of two occupations – waged employment, denoted by w, or self-employment, denoted by s. Human capital specific to each occupation is modeled by stock variables for each occupation.

Wage earnings, $y\left(\theta_w^t\right)$, are a function of the agent's stock of wage-specific human capital, θ_w^t, which can accumulate due to work experience. Let $\Phi_w\left(\theta_w^0, \cdot\right)$ denote the transmission of wage-specific human capital between the two periods, where the second argument of Φ_w records the agent's occupational choice in the first period. Then θ_w^0 denotes the initial endowment of wage-specific human capital, and $\theta_w^1 = \Phi_w\left(\theta_w^0, \cdot\right)$ denotes the stock of human capital in the second period. I assume that work experience has value, that is, that $\Phi_w\left(\theta_w^0, w\right) = \theta_w^1 > \theta_w^0$. For simplicity, $\theta_w^0 = \Phi_w\left(\theta_w^0, s\right)$ (i.e., if the agent does not acquire wage work experience then the stock of wage-specific human capital does not change). Note that this rules out the possibility that human capital relevant to wage employment can be acquired in self-employment. While it would be interesting to consider an extension that weakens this assumption, it would raise additional empirical challenges to identify cross-occupational-relevant human capital accumulation. Finally, I assume that Φ_w is increasing in its first argument.

The setup for self-employment is analogous. Self-employment earnings are influenced by the agent's stock of entrepreneurial human capital, θ_s^t. $\Phi_s\left(\theta_s^0, \cdot\right)$ denotes the transmission function, where $\theta_s^1 = \Phi_s\left(\theta_s^0, \cdot\right)$. Analogously, I assume that $\Phi_s\left(\theta_s^0, s\right) = \theta_s^1 > \theta_s^0$, that $\theta_s^0 = \Phi_s\left(\theta_s^0, w\right)$, and Φ_s is bounded.[10] The profit function is as follows, for $t = 0, 1$,

$$\begin{aligned} \pi\left(\theta_s^t, W^t, p^t, p_k^t, p_l^t\right) = \max_{l \geq 0} p^t f\left(\theta_s^t, k, l\right) - p_k^t\left(W - k\right) - p_l^t l \\ s.t.\ 0 \leq p_k^t k \leq \lambda W^t \end{aligned} \tag{1}$$

where p^t is the price of a single output in period t, k is capital, l is labor, p_k^t and p_l^t are their respective prices in period t, and f is an increasing, concave production function. I assume that the

[9] The early, seminal paper in this line is Jovanovic [1982]. Taveras [2010] carries out a calibration exercise on a similar model to show that a number of stylized facts that have been taken as evidence of credit constraints in prior literature can in fact be rationalized in a model of Bayesian learning about entrepreneurial skill if learning is sufficiently slow.

[10] Perhaps most important here is the second part of the assumption, which implies that entrepreneurial skills aren't acquired in wage employment.

the firm is a price-taker. The constraint set $k \in [0, \lambda W^t]$ is standard in the literature and captures credit constraints – the stock of physical capital employed in the enterprise may be constrained by own-funding constraints if there are frictions in credit markets and other financing sources are not available. That is, it may be that the optimal stock of capital, k^*, is strictly greater than λW^t, so that the firm is constrained from employing the optimal capital stock.

The timing of the model is as follows. The agent first draws human capital endowments θ_w^0 and θ_s^0 from the joint distribution η, with support on $\mathbb{R}^2$. This allows for arbitrary correlation between the two, which can be interpreted as capturing greater general ability in the individual, and can exacerbate selection effects as I will demonstrate below. These occupation-specific human capital stocks are known at the beginning of each of the two decision periods. Given these known human capital stocks, the agent makes a discrete occupational choice, between wage-employment and self-employment, w or s, in a forward-looking way in the first period. If the agent chooses self-employment, she makes a decision about the labor and capital inputs to the enterprise, k and l. The intertemporal connection between the two periods is given by the human capital transmission functions Φ_w and Φ_s as described above, along with the intertemporal savings problem. Denote the savings choice by x, where it must be that the value of x is less than the sum of wealth the agent opens the first period with, W^0, and earnings (y or π). The residual of the savings choice is consumption, which is evaluated in the strictly increasing, concave utility function U. At the end of the second period the agent is taken to consume all remaining wealth.

Formally, then the agent faces the following decision problem in the initial period, which is summarized by the value function V_0,

$$
\begin{aligned}
& V_0\left(\theta_w^0, \theta_s^0, W^0\right) && (2)\\
= \; & \max\Bigg\{ \max_{0\le x\le y+W^0} U\left[y\left(\theta_w^0\right) + W^0 - x\right] + \beta \max\left\{U\left[y\left(\theta_w^1\right) + W^1\right], \right. && (3)\\
& \left. U\left[\pi\left(\theta_s^0, W^1, p^1, p_k^1, p_l^1\right) + W^1\right]\right\}, \\
& \max_{0\le x\le \pi+W^0} U\left[\pi\left(\theta_s^0, W^0, p^0, p_k^0, p_l^0\right) + W^0 - x\right] + \beta \max\left\{U\left[y\left(\theta_w^0\right) + W^1\right], \right. \\
& \left. U\left[\pi\left(\theta_s^1, W^1, p^1, p_k^1, p_l^1\right) + W^1\right]\right\}\Bigg\} \\
= \; & \max\Bigg\{ \max_{0\le x\le y+W^0} U\left[y\left(\theta_w^0\right) + W^0 - x\right] + \beta \max\left\{U\left[y\left(\Phi_w\left(\theta_w^0, w\right)\right) + W^0 + x\right], \right. && (4)\\
& \left. U\left[\pi\left(\theta_s^0, W^0 + x, p^1, p_k^1, p_l^1\right) + W^0 + x\right]\right\}, \\
& \max_{0\le x\le \pi+W^0} U\left[\pi\left(\theta_s^0, W^0, p^0, p_k^0, p_l^0\right) + W^0 - x\right] + \beta \max\left\{U\left[y\left(\theta_w^0\right) + W^0 + x\right], \right. \\
& \left. U\left[\pi\left(\Phi_s\left(\theta_s^0, s\right), W^0 + x, p^1, p_k^1, p_l^1\right) + W^0 + x\right]\right\}\Bigg\}
\end{aligned}
$$

where the second equality illustrates the functional relationships that generate the final-period values

of the stock variables of occupational skill and wealth, and β is a discount factor in the $(0,1)$ interval. Namely, in the initial period the agent faces a discrete choice over the immediate occupational return given by y or π, and the discounted future return obtained from the same activity choice in the second period.

3.1 Basic Properties of the Model

The value function in equation (2) formalizes the dynamic incentives in the occupational choice problem. First, the individual faces an initial "selection" incentive, influenced both by the initial returns generated by the values of θ_s^0 and θ_w^0 (and possibly the effect of binding credit constraints on the physical capital decision), and prospective second-period returns due to savings and human capital accumulation. θ_s^0 and θ_w^0 might be accumulated prior to formally entering the workforce through familial effects, education, or other life experiences. All things equal we expect that individuals with a relatively higher stock of ability in a given occupation to be more likely to self-select into that occupation. Formally, the condition for selection into self-employment is as follows,

$$\begin{aligned} & \max_{0\leq x\leq \pi+W^0} U\left[\pi\left(\theta_s^0, W^0, p^0, p_k^0, p_l^0\right) + W^0 - x\right] + \beta V_1^s \\ \geq & \max_{0\leq x\leq y+W^0} U\left[y\left(\theta_w^0\right) + W^0 - x\right] + \beta V_1^w, \end{aligned} \tag{5}$$

where to save on notation V_1^s and V_1^w denote the future utility derived from choosing the optimal occupation in the second period, given human capital accumulated by the choices of s and w, respectively, in the initial period. The complementary condition captures the incentive for selection into waged employment. Equation (5) can be used to characterize the subsets of the parameter space under which selection into each occupation is optimal.

Self-selecting into a given occupation can lead to the acquisition of relevant human capital that further shifts the choice margin between the two occupations. That is, human capital accumulation can lead to lock-in, in a given occupation. If, for example, the individual chooses self-employment in the first period, this increases the value of θ_s^1, which increases the value of second-period profit $\pi\left(\theta_s^1, W^1, p^1, p_k^1, p_l^1\right)$ and hence increases the propensity to select into self-employment in the second period. In fact, dynamic incentives might even generate dynamic selection effects, under which individuals are incentivized to enter self-employment today even for a lower static return, under the anticipation of greater returns in the future (Foster and Rosenzweig [1995] capture similar intuition).

The human capital lock-in effect highlights the importance of the initial occupational choice, which is driven by the initial stocks of human capital, θ_s^0 and θ_w^0. If there is a reasonably high degree of correlation between the initial stocks of human capital, then it may be that the "highest-potential"

entrepreneurs do not enter self-employment at all, because the opportunity cost to self-employment is high based on wage earnings possibilities. This insight, first discussed semi-formally in the economics literature in Roy [1951], points out that occupational selection will be driven by the distribution of skills in the population and how they are compensated in equilibrium. It could be that most high-ability individuals tend to enter waged employment, and human capital lock-in further reinforces that choice. On the other hand, low-skill individuals might receive relatively lower returns in wage employment, particularly if low-skill labor supply is abundant. This is consistent with the massive cohort of low-skill, self-employed individuals in developing countries, most of whose enterprises have low returns and grow little.[11]

3.2 The Effects of Exogenous Shocks to Occupational Choice Incentives

The potential for human capital lock-in to prevent the highest-potential entrepreneurs from entering self-employment is suggestive of the empirical strategy that will be employed in this paper. I will seek a source of exogenous variation in selection incentives, orthogonal to individual ability, that leads relatively high-ability individuals to select into self-employment when they otherwise would not have done so. Such a shock can be due to any of the exogenous parameters of the model – to a price, to wealth, or to the earnings functions. The value function in equation (2) clarifies the effect of such changes, which are almost always unambiguous in the model. In this paper I will focus on an exogenous, negative shock to the wage employment earnings function, $y\left(\theta_w^0\right)$, though it is not problematic if the effects of the shock are transmitted through additional parameters of the model. Due to an exogenous event, which we can think of as occurring prior to period 0, the margin of choice will shift for a number of individuals, and they will have a much greater incentive to select into self-employment, as it becomes more likely that equation (5) will see a tilt in incentives toward self-employment.

3.3 Testable Predictions and Alternative Theories

A direct prediction of the theory is that entrepreneurial experience should lead to entrepreneurial persistence, even after accounting for opportunity costs. This is the human capital lock-in effect that was discussed above. As θ_s^0 increases to θ_s^1, the individual should be more likely to again engage in self-employment in period 1. This is true even if a negative wage shock caused an increase in

[11] This charactization is consistent with recent empirical evidence (e.g., Carter and Olinto [2003]; de Mel, McKenzie, and Woodruff [2008]; Banerjee, Duflo, Glennerster, and Kinnan [2009]; Karlan and Zinman [2010]). Demand for capital ends up being relatively stronger amongst wealthier or higher-ability individuals and hence individuals end up more responsive to positive financial shocks.

self-employment, and then the wage returns to its previous level. The empirical version of this prediction will be developed in Section 5.0.1. However, while the finding that individuals who obtain self-employment experience are more likely to remain self-employed is consistent with and strongly suggestive of a theory of entrepreneurial human capital accumulation, such a finding is still not a conclusive basis to argue that entrepreneurship-specific human capital accumulation is the primary factor driving enterprise dynamics. There are at least two alternative theories that generate a similar prediction, which in contrast take the stock of entrepreneurial skill as *fixed* and generate dynamic effects through other channels.

In Jovanovic [1982], individuals are endowed with a fixed stock of entrepreneurial skill, which they are uncertain about and have prior beliefs over. In the context of the model developed herein, we can think of this as an entrepreneurial skill parameter θ_s that doesn't vary over time, but determines the distribution of stochastic realizations of the production function.[12] The individual holds subjective beliefs μ_{θ_s} over the distribution of θ_s, which is initially drawn from a normal distribution with known mean and variance. Since the individual does not know the exact value of her own θ_s, the initial belief is taken as the mean of the distribution μ_{θ_s}. Over time, as the firm operates, the agent draws observations on a stochastic production process, which allow for inferences on θ_s, with updating of beliefs through a standard Bayesian learning process. Hence there is a co-movement of beliefs and firm size – in expectation good entrepreneurs grow their firms as their beliefs about own-ability move upwards, while bad entrepreneurs shrink and eventually exit.

In Buera [2009] and related models, skill is known but credit market constraints might prevent the optimal physical capital level from being attained, at least in the short run. Individuals know their entrepreneurial skill level, and indeed that knowledge may exactly induce them to save more ex ante in order to eventually be able to self-fund the startup or growth of an enterprise. In the context of the model herein, we can think of this as a case where the function $\Phi_s(\cdot,\cdot)$ is a constant function. For the model to be empirically relevant it is important that the credit constraint actually binds for a significant proportion of the population. Similarly to Jovanovic [1982], the theory predicts that physical capital increases over time for good entrepreneurs, as it converges to the level most compatible with the endowment of skill.

Hence both models allow for the possibility that a significant number of individuals who enter self-employment will be persistent and see an increase in inputs and earnings over time. They suggest that a significant number of enterprises will enter the market at a different scale from their long-run optimal scale, and that successful firms will converge to the long-run optimal size as dictated by the fixed stock of entrepreneurial skill. Of course, one can quickly see that the theory of entrepreneurial human capital accumulation will also predict increases in capital and labor inputs over time, to optimally

[12] In fact, in Jovanovic [1982] θ_s is a parameter that determines the distribution of shocks to the cost function.

complement the stock of entrepreneurial skill. However, what is critical is that the alternative theories suggest that increases in returns should be *input-driven,* in terms of labor and capital inputs. The theories do not allow for residual increases in profitability due to increases in the entrepreneurial and managerial abilities and business capital of the individual running the firm.

An additional test would then be to study the relationship between earnings increases of the firm and increases in the size of the capital and labor stock of the firm. The testable prediction disentangling a model with dynamically-accumulating entrepreneurial human capital from the other theories would be the finding that the variation in earnings could not be explained by dynamic changes in inputs alone. Technically this requires some assumptions about the revenue function of firms, in order to discipline the relationship between inputs and revenues.

A final notable implication regards the *rate* of purported entrepreneurial human capital accumulation amongst entrants. It seems reasonable to posit that the learning function takes a concave shape, with diminishing returns to learning as more human capital is accumulated. This would mean that, all things equal, brand new entrants (those who had not previously run an enterprise) should learn at the highest rate, and hence be subject to the largest change in earnings and occupational choice incentives. We will also look for support for this final testable implication.

4 Design of the Study and Preliminary Evidence

4.1 Data

My primary dataset is the Indonesia Family Life Survey (IFLS).[13] The data were collected as a household panel survey in Indonesia, with data collection rounds in 1993, 1997-98, 2000-01 and 2007-08. The 1997-98 round directly proceeded the crisis. For the intervening years when the survey is not fielded, significant retrospective data are collected in the subsequent round. The dataset was designed to be representative of 83% of the Indonesian population in 1993, covering 13 of the higher-population provinces generally in the western parts of the country, with over-sampling of urban locations and locations outside Java island, the main economic hub. Data were collected at the individual, household, and community level, and these three sources can be matched together. More details on relevant parts of the dataset, including for enterprise activity, will be discussed in more detail below.

The original 1993 round of the survey (IFLS1) surveyed 7224 households. Subsequent rounds

[13] Various organizations and researchers have been involved in designing, collecting and funding the IFLS. For more details, see Strauss, Witoelar, Sikoki, and Wattie [2009], Strauss, Beegle, Sikoki, Dwiyanto, Herawati, and Witoelar [2004], Frankenberg and Thomas [2000], and Frankenberg and Karoly [1995].

have involved re-sampling the original households, and then sampling all split-offs from the original households. Attrition has been relatively minor, at less than 10% between rounds, and overall 87.6% of the original households appear in all four rounds. Table 1 presents the number of individuals,[14] households, household enterprises and communities appearing in each round of the survey. We see that the sample expands in each subsequent round, as splits from the original households are tracked and surveyed. In addition, the proportion of household members directly interviewed also increases across rounds.

There is significant geographic and size variation amongst the enterprises.[15] Though the largest firm representations are from Java, the economic and population center of the country, the bias is not overwhelming and a significant proportion of firms are observed from all of the main survey provinces. This is true even if we focus on firms with a relatively larger capital stock, above $1000 US (converted from Indonesian rupiah at the going exchange rate in a given survey year). It is notable that the slightly larger proportion of firms seems to be in rural areas. This fits with Liedholm and Mead [1999] and may be due to the fact that smaller firms are more likely to service demand in more remote areas. Also, we see that the sample contains a significant number of firms exceeding the sizes observed in the vast majority of studies on micro and small enterprises from developing countries, while firm-level surveys looking at such firms generally have little information on the primary entrepreneur. Given that conversion to US purchasing power parity implies a multiple of about 12, there are hundreds of enterprises with more than $25,000 US PPP equivalent in capital, and dozens with 10, 15 or more workers.

Table 3 presents a summary of a number of community-level measures of market churning that will be useful in the background of the later analysis, as these variables are used as exogenous sources of variation in the individual propensity to enter and remain in self-employment.

4.2 Preliminary Evidence

In Figure 1 I non-parametrically plot experience-earnings (net profit) profiles across these three qualitative categories, using a Lowess tri-cube smoother. There we see that while all three groups enjoy an increase in earnings on average, the rate of increase is substantially higher for those running the enterprise we would expect to be most complex: firms with hired, wage workers. This bifurcation in returns is suggestive of the select group of individuals running more complex enterprises "pulling

[14] Both adults and children (defined as those under age 15 at the time of the survey) are surveyed, though the childrens' module is less extensive.

[15] The distribution of enterprises is less even if we stratify by industry–the largest proportions of enterprises by far are in the sectors of restaurant/food, and sales:non-food, at around 30% each. The next two largest sectors are food processing, and services:transport.

away" from the much larger group of individuals running enterprises in the other two categories. We would expect that significantly greater returns would enable significantly greater capital accumulation.

In Table 2 I present summary statistics on the smaller population of individuals who enter self-employment during the financial crisis, a smaller sample. There are 684 such individuals who are eligible for the study due to entry during 1997, and 1355 eligible due to entry in 1998. We see that they are highly likely to be married, often quite well-educated, and more likely to be male. They also appear relatively younger, which could be a reflection of the role in seniority in worker separations during the crisis.

5 Identification Strategies and Empirical Specifications

The identification of EHC raises empirical challenges due to the selection processes highlighted in the model. The ideal experiment would randomly assign EHC to individuals, orthogonally to all other characteristics, and then observe the resulting enterprise performance trajectories. Clearly such an experiment would be infeasible for a number of reasons, including endogenous enterprise survival, and difficulties in assigning EHC. However, individuals with greater *ex ante* (unobserved) entrepreneurial ability are more likely to select into self-employment, and hence accumulate greater entrepreneurial human capital. Hence higher-ability self-employed individuals are likely to have better entrepreneurial performance (i.e., higher enterprise returns), while simultaneously having greater accumulated experience, due to endogenous survival effects.

Since the ideal experiment is not feasible in practice, I exploit a source of exogenous and unanticipated assignment into self-employment (experience) due to the 1997-98 Financial Crisis. Here the primary "treatment" group of interest is the subset of individuals that enter self-employment during 1998, the main year in which the effects of the crisis were felt in Indonesia. In particular, the interest is in individuals who were 'pushed' into self-employment, who would not have otherwise entered, which provides a source of a counterfactual to consider the effects of the quasi-random assignment of EHC.

Previous analyses of the effects of the crisis have shown that the crisis did not cause a significant drop in overall employment; however, it caused a significant shift in real wages, in some cases up to 40%, with effects particularly concentrated on relatively higher-earning, formal-sector wage workers (Thomas, Beegle, and Frankenberg [2000]). This exogenous shock is particularly useful for the purposes of this study, because it means that a significant number of relatively higher-ability individuals were 'pushed' into self-employment. Hence this natural experiment is quite appealing to test the

theory of EHC, because other sources of exogenous shocks such as rainfall might be expected to be concentrated on lower-income individuals who might have fewer alternatives to self-employment.

Based on this intuition, I construct two main tests of the interpretation of EHC as a natural experiment that assigns individuals to acquire entrepreneurial experience. First, I look at self-employment persistence. The theoretical model predicts that individuals who acquire human capital in a particular occupation should, all things equal, be more likely to remain in that occupation. In testing this implication I deal with the assumption of "all things equal" potentially not holding by using a number of regression controls, including in particular an out-of-sample estimate of the counterfactual wage that self-employed individuals would earn in wage employment. Since dynamic occupational persistence can be explained by a number of theories outside of EHC accumulation, I secondly look at earnings dynamics, to generate further evidence consistent with the proposed theory.

Building on this evidence, I provide evidence on the causal effect of entrepreneurial learning-by-doing on earnings. The empirical challenge that arises is that, of course, the crisis is not a perfect natural experiment. It does not necessarily randomize selection into self-employment (and subsequent acquisition of experience) orthogonally to unobserved ability, in particular. Those cohorts that enter self-employment, even during the crisis period, presumably include at least two groups: (i) those who enter self-employment as a survival response to the shock (due to having lost their job, etc.), or the 'pushed' group of interest, and (ii) those who enter self-employment voluntarily (perhaps because the disequilibrium process highlights a new profit-making opportunity), or because they were already planning to enter self-employment independently from the crisis. I attempt to control for endogenous selection into self-employment in 1998 through a selection-on-observables-type strategy, which is plausibly exogenous to individual-level EHC.

Details behind these approaches are discussed in the remainder of this section, and estimation results are then presented in Section 6.1.

5.0.1 Self-employment Persistence

An important implication of the theoretical model is that the accumulation of EHC changes the occupational choice incentives of the individual. If in period 1 the individual chooses self-employment, s, perhaps due to a shock to the opportunity cost to self-employment (the wage $y\left(\theta_w^0\right)$), then the human capital accumulation function, $\Phi_s\left(\theta_s^0, s\right)$, implies that the stock of EHC increases from θ_s^0 to $\theta_s^1 > \theta_s^0$. Even if the opportunity cost of self-employment returns to near its previous level, the agent is more likely to find it optimal to remain in self-employment in subsequent periods. This trade-off is formalized in the model, in particular where we see that the second-period decision involves the

static maximization problem,

$$\max \left\{ U \left[y \left(\theta_w^1 \right) + W^1 \right], U \left[\pi \left(\theta_s^1, W^1, p^1, p_k^1, p_l^1 \right) + W^1 \right] \right\}. \tag{6}$$

Of course the outcome is not deterministic – since the choice is discrete it may be that the choice margin moves but still not enough to induce the agent to remain in self-employment once the wage recovers. However, across the population distribution we might expect to observe an effect.

Just looking at raw numbers, we see that 684 individuals newly shift into self-employment in 1998, about a 10% increase in the number of self-employed individuals. By the year 2000, 587 of these individuals are still self-employed (about 85%), even though the economy has already shown significant recovery from the crisis event. Even by the year 2008, about 78% remain in self-employment. This comes in stark contrast to the comparable figure from other the 10 years of the survey from which a 9-year persistence rate can be calculated, which averages 46%. However, these raw indications are subject to some key challenges in terms of identification. I discuss how I deal with these challenges in what follows.

Firstly, the descriptive evidence on self-employment persistence does not control for the expected wage, the opportunity cost to self-employment. It could be the case that wages don't recover for the types of individuals who enter self-employment during the crisis, and hence in fact the opportunity cost to self-employment remains low. In addition, as has been already discussed, the self-employment entry decision can be driven by unobservables. In order to account for these concerns, I carry out the following estimation procedure to attempt to provide more convincing evidence for the robustness of occupational persistence.

The intuition behind the procedure is to construct the (unobserved) counterfactual probability of being self-employed in absence of having entered self-employment during the crisis, and then compare that to two constructs of the realized propensity to be self-employed: (1) the empirical realization of self-employment propensity (the simple frequentist estimate), and (2) an estimated probit model on ex post occupational choice outcomes in the sample of individuals who enter self-employment during the crisis. Hence it involves a within comparison of predicted entrepreneurial propensity based on pre-crisis choices, with ex post realized outcomes. I will then argue that results of sufficient magnitude overcome other explanations for the self-employment persistence such as, e.g., labor market frictions preventing re-integration into the formal wage sector.

The procedure works as follows. First, I empirically capture the occupational choice decision rule of individuals by estimating a probit self-employment selection equation of the form,

$$\Pr \left(y_{it} = 1 | \mathbf{x}_{it} \right) = G \left(\mathbf{x}_{it} \boldsymbol{\beta} + \boldsymbol{\gamma} \boldsymbol{\omega}_{it} \right), \tag{7}$$

where y represents the discrete occupational choice ($y_{it} = 1$ denotes self-employment, and $y_{it} = 0$ denotes wage employment), G is the standard normal density, $\mathbf{x}_{it}$ is a vector of regression controls such as age and age-squared, education (in years) and marital status, $\boldsymbol{\omega}_{it}$ represents the wage, and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ represent regression coefficients.

I estimate the above model using two definitions of the population. First, just on the sub-sample of individuals who enter self-employment during the crisis, and secondly on the whole population. The former more directly captures the choice function of the specific individuals involved, though it might underestimate entrepreneurial propensity since these individuals are less likely to be self-employed pre-crisis. The latter better captures the determinants of entrepreneurial selection in the population, though it might induce estimates that are less applicable to the particular crisis-entrant sample.

Second, having used this model to estimate the occupational choice function, I then construct projected occupational selection propensities, $\hat{G}_{it}$ out of equation (7), by predicting out of sample using the estimators $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$. Since the wage, $\boldsymbol{\omega}_{it}$ in equation (7), is not observed once individuals have entered self-employment, I employ the following wage equation in order to calculate the individual-specific projected wage, $\hat{\boldsymbol{\omega}}_{it}$, as a measure of the opportunity cost of self-employment,

$$\boldsymbol{\omega}_{it} = \mathbf{x}_{it}\boldsymbol{\delta} + c_i + year_t + \varepsilon_{it}, \tag{8}$$

where $\mathbf{x}_{it}$ is a vector of regression controls such as age and age-squared, education (in years) and marital status, c_i is an individual-specific fixed effect term, and $year_t$ is a year effect.[16] I use a bootstrap approach to deal with the issue of using projected regressors as explanatory variables in a subsequent regression.

The out-of-sample prediction of self-employment propensity, $\hat{G}_{it}$, gives a counterfactual measure of self-employment propensity. I generate out-of-sample predictions of $\hat{G}_{it}$ from two different data samples, which I denote $P^{sub}_{Pre \rightarrow Post}$ and $P^{full}_{Pre \rightarrow Post}$, respectively. I denote the mean of the distribution of values of individual-specific self-employment propensities based on ex ante data only from the *subsample of crisis-period entrants* by $P^{sub}_{Pre \rightarrow Post}$. I denote the same object, estimated on *full ex ante population* data, by $P^{full}_{Pre \rightarrow Post}$.

Third, I construct ex post measures of self-employment propensity, from actual realizations in the data. I denote by $P^{sub}_{Post\ freq}$ the empirical realization of self-employment propensity (the simple frequentist estimate), and by $P^{sub}_{Post\ prob}$ occupational choice propensity estimates generated from a probit model on ex post occupational choice outcomes in the sample of individuals who enter self-employment during the crisis.

[16] I do not include time-variant, location-specific variables as controls, since geographic identifiers are not always available for each observation, meaning sample size would be noticeably reduced.

I then test whether there is a statistically significant difference in the propensity to be self-employed, comparing the constructed counterfactuals, $P^{sub}_{Pre \rightarrow Post}$ and $P^{full}_{Pre \rightarrow Post}$ to the ex post realizations, $P^{sub}_{Post\ freq}$ and $P^{sub}_{Post\ prob}$. I apply t-tests to the differences in the means of the two distributions.

5.0.2 The Dynamics of Self-Employment Returns

The analysis of persistence provides a convincing source of evidence on the propensity to remain self-employed. Yet, it does not rule out some alternative hypotheses outside of the endogenous accumulation of EHC. The main competing theories of entrepreneurial dynamics take entrepreneurial ability as fixed, and then generate dynamics from learning about own-ability Jovanovic [1982], saving, or the like. To disentangle the proposed theory of EHC accumulation from a Jovanovic-type story, I study enterprise earnings dynamics.

In Jovanovic's model, individuals persist in self-employment because they turn out to be the 'good' entrepreneurs, through getting earnings draws and learning about own ability. In such a model we should not see entrepreneurial returns increase greatly relative to the overall economy, because optimal entrepreneurial inputs are available immediately at enterprise startup. Hence I study the dynamics of enterprise earnings and how they increase relative to the growth of the overall economy and counterfactual wages.

Selection-corrected earnings dynamics The final piece of evidence on earnings dynamics comes from taking years spent running an enterprise as a measure of learning-by-doing and entrepreneurial human capital acquisition. This is analogous to the literature on education as a source of human capital. Similarly to that literature, the main empirical problem in deriving causal estimates of the effects of human capital acquisition is one of selection: individuals are not randomly assigned to acquire entrepreneurial experience. I take the crisis to provide quasi-experimental variation in the incentives to enter self-employment and hence acquire entrepreneurial experience.

I calculate years of experience running enterprises in three different size categories – no employees, only household/unpaid employees in the enterprise, or those which hire permanent wage workers for an explicit wage. I then use an adaptation of the Heckman selection procedure to study the selection-corrected relationship between the experience measures and self-employment earnings (net profit).

The traditional Heckman model involved running a first-stage selection equation, then using it to generate an individual-level estimate of the propensity to select into one of the selection options, which is then fed into the second-stage equation as the inverse Mills ratio. I follow this approach, inserting an estimate of the propensity to enter self-employment which has already been presented

above in Section 5.0.1, in equation (7). I take the appropriate version of $\hat{G}_{it}$ to give me the individual-level occupational selection propensity, then use it as a control in an earnings experience regression. As already discussed above, the first-stage selection model incorporates variables reasonably excluded in the second-stage earnings equation – primarily location-level measures of occupational churning. Hence the requirement of at least one non-intersection between the first- and second-stage equations in a traditional Heckman setup is satisfied. This approach accounts for individual-level variation in the propensity to enter self-employment in a given period, based on observables.

The second-stage earnings equation is given as follows:

$$y_i = \beta_0 + \exp_i \boldsymbol{\beta} + x_i \boldsymbol{\gamma} + \hat{G}_i \boldsymbol{\delta} + \boldsymbol{\varepsilon}_i, \tag{9}$$

where y_{it} represents reported self-employment earnings by individual i, β_0 is a constant, $\exp_i$ is a vector of individual-specific entrepreneurial experience counts, x_i is a vector of other controls (age, age^2, gender, education (in years), marital status), and $\hat{G}_i$ is the projected occupational selection value.

Given that this procedure introduces a generated regressor in the second-stage earnings equation through $\hat{G}_i$, in the second stage estimates I use a bootstrap procedure with 50 replications, to account for a potential non-standard error distribution rather than imposing normality on the model.

6 Estimation Results: Self-employment Persistence and Returns

In this section I present the empirical results on self-employment persistence, and self-employment earnings dynamics, respectively.

6.1 Self-employment Persistence

The results from this part of the analysis are summarized in Tables 4 to 6.

I present the results of the fixed effects wage equation analysis in Table 4. The main goal of this equation is explanatory power, and that seems to be achieved with an R^2 of 0.49. The age effect is notable in implying a convex function, though the first-order coefficient is negative but not statistically-significant. In general Mincer equations will generate a concave age effect. Otherwise we find that the remaining regressors are almost always highly statistically-significant with reasonable coefficients.

Looking at entrepreneurial persistence, I report on analysis looking separately at the group of individuals who enter self-employment during 1998 (which might be affected by the onset of the crisis), and those who enter self-employment during 1999 (whose employment incentives would be expected to be heavily affected by the brunt of the crisis), in Tables 5 and 6, respectively. I find that the propensity of the individuals who enter during the crisis, which can be reasonably argued to be dominated by those involuntarily forced into self-employment, to remain in self-employment is remarkably high, even after controlling for the opportunity cost of self-employment, the expected wage. All changes in propensity are strongly statistically significant, by a standard t-test.

Individuals who are self-employed during the crisis are very likely to be self-employed even after the crisis – about a tripling of the propensity to be self-employed for the 1998 entrants, and anywhere from a four to ten times increase for those who enter in 1999. As we look at years further and further from the crisis, up until 2008, the propensity to remain in self-employment remains remarkably strong. Namely, even after accounting for the expected wage, and hence the recovery of the economy, we still see a very strong increase in propensity to remain self-employed. I take this to suggest that the individuals who involuntarily enter self-employment during the crisis manage to accumulate a significant enough stock of EHC from that experience that they become much more likely to subsequently engage in self-employment. I argue that this increased propensity is far greater than would be predicted by any reasonable model that assumes away the endogenous accumulation of EHC. In particular, the effect seems to be so qualitatively large as to exceed any reasonable frictions that might inhibit back into the wage sector, such as job search frictions.

6.2 The Dynamics of Self-Employment Returns

If we look at the raw numbers, we see that in the year in which the main crisis cohort enters self-employment, 1999, their self-employment earnings are about 9% lower than the counterfactual expected wage. Note that this is perfectly reasonable in a model in which (i) individuals are unexpected forced to enter self-employment (due to the crisis), and/or (ii) they anticipate dynamic increases in earnings over time. By 10 years later the situation has flipped quite strongly – expected earnings are now 16% higher than the counterfactual expected wage. While the latter figure is biased somewhat by the natural attrition of some lower-performing entrepreneurs, the bias is limited by the low attrition that has already been discussed in this cohort.

This effect seems large in terms of levels, also. By the year 2000, 1998 entrants see a 40% increase in profitability, while 1999 entrants see a 20% increase. This is substantial, and far exceeds the growth rate of the economy (as a control for time trend). For example, we see only about an 8% increase in expected wage, which provides a reasonable and context-relevant baseline comparison to control

for economic growth overall. This trajectory in returns points to a human capital accumulation dynamic. In addition, it provides further evidence against labor market frictions in explaining the lack of switching – switching costs would have to be on the order of 20% of yearly income to justify not switching back into wage employment.

6.3 Selection-corrected earnings dynamics

The final piece of evidence on earnings dynamics comes from taking years spent running an enterprise as a measure of learning-by-doing and entrepreneurial human capital acquisition. As discussed, I employ a version of the Heckman selection model. As also noted, part of the first-stage analysis is taken from previous work on occupational persistence, and hence I do not present those first-stage results here, rather just focusing on the second-stage earnings equation.

Tables 7 and 8 provide final-stage selection-corrected evidence on returns, using entrants from the years 1998 and 1999, respectively. What we first notice is that selection bias, at least according to controls based on observables, does not appear to be an important problem, as the estimated coefficient on the inverse Mills ratio is not statistically distinguishable from zero in either regression.

Looking at the coefficients on the experience variables, we see that the shape of returns in experience is intuitive, following a concave shape for all three types, with the exception of the single proprietor enterprises amongst 1999 entrants, with easily the highest returns for those running enterprises in the greater complexity category. Among the 1998 entrants, the estimated learning effect is positive for all three types, though seemingly less persistent for those running the simplest enterprises, where the positive learning effect diminishes after just over 4 years. By contrast, the learning effect persists for around 6 years for the other types. Keeping in mind that median enterprise experience is around 5-6 years in the population, it seems safe to conclude that there is a positive learning effect. Many of the coefficients for the 1998 entrants are not statistically significant. While it is tempting to rationalize this based on low sample size, we get much more precisely-estimated effects from the smaller 1999 entrant cohort.

Looking at the results from the 1999 cohort, there is interestingly a convex estimated experience effect for individuals running the simplest enterprises, which actually implies negative returns to experience for the first five plus years running an enterprise. While the initial estimated learning effect is positive in the other two enterprise types, it is dissipates quickly.

These results emerge after controlling for endogenous selection into self-employment, again with variables plausibly exogenous to individual EHC endowments. In other words, this evidence is about as close as we could reasonably expect to get to exogenously assigning experience to individuals.

7 Alternative Explanations and Further Evidence

While the analysis above is derived from the two strongest sources of evidence to support the interpretation of the natural experiment – occupational persistence and earnings dynamics – the evidence does not account for all alternative explanations. This section presents tests meant to account for and cast doubt on alternative explanations for the results.

7.1 Capital Stock Lock-in

One possible alternative explanation for enterprise persistence, other than the posited story of entrepreneurial human capital acquisition, is capital stock lock-in. Namely, that individuals who entered during the crisis took on greater sunk costs in their capital stock investments, which they might have been reluctant to abandon as the economy recovered. While this seems somewhat implausible as the crisis period was a time of great uncertainty that saw greater levels of investment from inexperienced entrepreneurs (both of which should lead to less significant investment), I carry out a test.

I test this by looking at capital stocks held by individuals who enter self-employment during the crisis period, and those who had entered at other times. What we would expect is that if capital stock lock-in were to explain enterprise persistence, the quantities of startup capital should be larger for firms that entered during the crisis years. Since data on startup capital is only provided in IFLS4, I am forced to focus on IFLS4 as the source of data. If we look at average startup capital in the full sample of firms, it is 6722876 Rph., while for firms which started up in either 1998 or 1999 (following the main crisis years), the average starting capital is 5027424 Rph. While both figures are somewhat biased due to survival (of firms still active during the 2007-08 survey round), since they both are subject to the same type of bias this is less of a concern.

7.2 Optimal Industry Selection

Another possible explanation for the positive earnings dynamic among survivors after the crisis years is that the individuals who start such enterprises during the crisis might have optimally selected into higher-growth industries (due to skill or luck, or both). Hence their earnings increases would be better explained by the effects or riding a wave during the period of opportunities that a crisis brings about. The initial suggestion seems implausible in light of the fact that the crisis was a fairly long-lasting disruptive event, and there continued to be economic and political changes well beyond the initial onset which would have made it difficult for early entrants into self-employment to parse out the best opportunities.

7.3 Changes in Inputs

Another potential explanation for the positive earnings dynamic is that it is driven by inputs. Namely, whether it is because individuals don't know their optimal input mix initially (and need to learn) or because of market constraints preventing initial access to the optimal input mix for new entrants, firms will increase input usage over time, and hence naturally increase returns.

However, this assertion does not fit with what we know about changes in capital stock and labor stock amongst firms. Tables 9 to 10 summarize the evidence on firms in the IFLS. In particular, we focus on observed changes in enterprise size from startup to present, the last 2 rows of the tables. The tables record transitions from startup size to current size for all of the enterprises in the sample that are operating in 2008.

We see that the propensity of household enterprises to significantly change in size is quite small, whether size is measured in terms of physical capital or labor. Only 14% of firms show any growth at all in labor, and for most the growth is minimal. Even as of the 95th percentile of the distribution, firms show no change in labor stock. The average change in labor employment is actually a small decrease. Looking at the data on capital stock, we again see minimal changes in firm size; even as of the 75th percentile of the capital stock growth distribution, we only have about a $135 USD change in physical capital since startup.

Taken together, this evidence indicates that firm growth is not driven by increases in inputs alone.

8 Conclusion

In this paper I develop and test a microeconomic theory of entrepreneurial human capital (EHC) accumulation. The key channel for acquiring EHC is through direct learning-by-doing. The theory is tested through exploiting a natural experiment based on the 1997-98 Indonesian financial crisis, which provides a source of exogenous assignment into entrepreneurial activity. This is useful because it pins down some confounding factors in selection into self-employment. The setting is also particularly suitable because we observe a large cohort of formal-sector workers exogenously forced into self-employment, in the more-stable informal sector. Consistent with the theory of entrepreneurial human capital accumulation, entrepreneurial activity is remarkably and robustly persistent. Even after controls for opportunity costs the propensity to be self-employed amongst this cohort increase by 2 to 9 times. The selection-corrected dynamic increase in returns to self-employment exceeds what could be reasonably expected in the absence of human capital accumulation. Taken together, these results suggest the importance of modeling entrepreneurial dynamics in a way that incorporates the role of endogenous human capital accumulation.

These results have a number of implications for policies regarding entrepreneurship promotion in developing countries. First, they highlight the importance of the accumulation of entrepreneurial human capital in enterprise outcomes. The policy implication from a model in which ability is fixed is that it is the financier's job to identify the ex ante higher-ability types as soon as possible and provide them with the full complement of financing that is proportional to their stock of entrepreneurial skill. A theory of dynamic entrepreneurial learning, however, suggests a more incrementalist approach with greater attention to timing, mixing financing provision with other skill-building services.

Second, the results pertain to institutions for the transfer of entrepreneurial human capital. In most countries the primary institution for the formation of skills for the waged-sector is formal education, which can last twelve or more years. While some writers, notably Schultz, have suggested that education might be an important venue for the formation of entrepreneurial skill, such a hypothesis is not well supported by the evidence in this paper. Instead, the results suggest that entrepreneurial skills are more specific and require more focused and sustained exposure to enterprise activity itself.[17] Hence this suggests the potential for specialized institutions for the transfer of entrepreneurial human capital. In most developing countries, the existing institution seems to be the family unit, at least those households in which the parents have a significant stock of entrepreneurial human capital that can be transferred to their children.

There have attempts at various forms of entrepreneurial training, including recent tests in the economics literature based on RCT designs, but based on the results in this paper it is not so surprising that the results from short-term training have been mixed at best. While many of the existing programs are focused on transferring low-level entrepreneurial skills (keeping records, basics of managing finances, etc.), it seems that high-level entrepreneurial skills (sales, marketing, risk judgment, product development, etc.) may be significantly more important, particularly for growth-oriented firms. It may be that a more intensive, sustained mix of direct experience and perhaps mentorship from more experienced and successful entrepreneurs is needed to enable the emergence of higher-potential entrepreneurs and the transfer of high-level entrepreneurial skills.

This paper also raises a number of questions for future research. The most obvious one regards the identification of entrepreneurial human capital and its various components. What are the most important high-level entrepreneurial skills? Are they complementary to each other, or are certain skills critical at certain stages? How can such skills be effectively transmitted? A key identification challenge faced by this paper is that many of the results could be explained not by entrepreneurial human capital accumulation that is internal to the entrepreneur, but rather an external reputation-building process amongst customers and other business partners. Of course, reputation is heavily

[17] This is not to suggest that education is not useful in general, particularly for pushing up the overall level of human capital in the population. However, the evidence herein, based on within-population variation in education and EHC, suggests that EHC is a more important *relative* factor in enterprise outcomes.

entangled with the underlying ability and performance of the entrepreneur in question. Hence future research might employ research strategies better suited to teasing out these internal and external effects.

References

Angrist, J. (1990): "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *The American Economic Review*, 80(3), 313–336.

Banerjee, A., E. Duflo, R. Glennerster, and C. Kinnan (2009): "The Miracle of Microfinance? Evidence from a Randomized Evaluation," MIT working paper.

Barrett, C., M. Bezuneh, D. C. Clay, and T. Reardon (2005): "Heterogeneous Constraints, Incentives and Income Diversification Strategies in Rural Africa," *Quarterly Journal of International Agriculture*, 44(1), 37–60.

Becker, G. (1964): *Human Capital*. Columbia University Press, New York, NY.

Behrman, J., and M. Rosenzweig (1999): "'Ability' Biases in Schooling Returns and Twins: A Test and New Estimates," *Economics of Education Review*, 18(2), 159–167.

Bloom, N., A. Mahajan, D. McKenzie, and J. Roberts (2010): "Why Do Firms in Developing Countries Have Low Productivity," *American Economic Review Papers and Proceedings*, 2, 619–623.

Bruhn, M., D. Karlan, and A. Schoar (2010): "What Capital is Missing in Developing Countries?," *American Economic Review Papers and Proceedings*, 2, 629–633.

Buera, F. (2009): "A Dynamic Model of Entrepreneurship with Borrowing Constraints: Theory and Evidence," *Annals of Finance*, 5(3), 443–464.

Cagetti, M., and M. D. Nardi (2006): "Entrepreneurship, Frictions and Wealth," *Journal of Political Economy*, 114(5), 835–870.

Card, D. (1999): *The Causal Effect of Education on Earnings*vol. 3A. Elsevier.

Carter, M., and P. Olinto (2003): "Getting Institutions Right for Whom? Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment," *American Journal of Agricultural Economics*, 85(1), 173–186.

CIA (2012): "C.I.A. Factbook 2012," Available at: [https://www.cia.gov/library/publications/the-world-factbook/].

de Mel, S., D. McKenzie, and C. Woodruff (2008): "Returns to Capital: Results from a Randomized Experiment," *Quarterly Journal of Economics*, 123(4), 1329–1372.

——— (2010): *Who are the Microenterprise Owners? Evidence from Sri Lanka on Tokman V. De Soto*pp. 63–88. The University of Chicago Press.

Foster, A. D., and M. R. Rosenzweig (1995): "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103(6), 1176–1209.

Frankenberg, E., and L. Karoly (1995): "The 1993 Indonesia Family Life Survey: Overview and Field Report," Discussion paper, RAND Corporation, Santa Monica, CA.

Frankenberg, E., and D. Thomas (2000): "The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and," Discussion paper, RAND Corporation, Santa Monica, CA, DRU-2238/1-NIA/NICHD.

Iyigun, M., and A. Owen (1998): "Risk, Entrepreneurship, and Human-Capital Accumulation," *American Economic Review Papers and Proceedings*, 88(2), 454–457.

Jovanovic, B. (1982): "Selection and the Evolution of Industry," *Econometrica*, 50(3), 649–670.

Karlan, D., and J. Zinman (2010): "Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate Impacts in Manila," Working paper.

Klein, P., and M. Cook (2006): "T.W. Schultz and the Human-Capital Approach to Entrepreneurship," *Review of Agricultural Economics*, 28(3), 344–350.

Liedholm, C., and D. Mead (1999): *Small Enterprises and Economic Development: The Dynamics of Micro and Small Enterprises*. Routledge, New York, NY.

Otani, K. (1996): "A Human Capital Approach to Entrepreneurial Capacity," *Economica*, 63(250), 273–289.

Poppele, J., S. Sumarto, and L. Pritchett (1999): "Social Impacts of the Indonesian Crisis: New Data and Policy Implications," Discussion paper, World Bank, Jakarta, Indonesia, SMERU report.

Porta, R. L., and A. Shleifer (2008): "The Unofficial Economy and Economic Development," *Brookings Papers on Economic Activity*.

Roy, A. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers (New Series)*, 3, 135–146.

Schoar, A. (2010): *The Divide Between Subsistence and Transformational Entrepreneurship*vol. 10. The University of Chicago Press, Chicago, IL.

Schultz, T. (1980): "Investment in Entrepreneurial Ability," *The Scandanavian Journal of Economics*, 82(4), 437–448.

Shane, S. (2003): *A General Theory of Entrepreneurship: The Individual-Opportunity Nexus*. Edward Elgar, Cheltenham.

Smith, J., D. Thomas, E. Frankenberg, K. Beegle, , and G. Teruel (2002): "Wages, Employment and Economic Shocks: Evidence from Indonesia," *Journal of Population Economics*, 15, 161–193.

Strauss, J., K. Beegle, B. Sikoki, A. Dwiyanto, Y. Herawati, and F. Witoelar (2004): "The Third Wave of the Indonesia Family Life Survey (IFLS)," Discussion paper, RAND Corporation, Santa Monica, CA, DRU-2238/1-NIA/NICHD.

Strauss, J., F. Witoelar, B. Sikoki, and A. M. Wattie (2009): "The Fourth Wave of the Indonesian Family Life Survey (IFLS4): Overview and Field Report," Discussion paper, WR-675/1-NIA/NICHD.

Taveras, C. (2010): "Entrepreneurship, Learning and Wealth," MIT mimeo.

Thomas, D., K. Beegle, and E. Frankenberg (2000): "Labor Market Transitions of Men and Women During an Economic Crisis: Evidence from Indonesia," RAND Labor and Population Program working paper series 00-11.

A Appendix

A.1 Figures

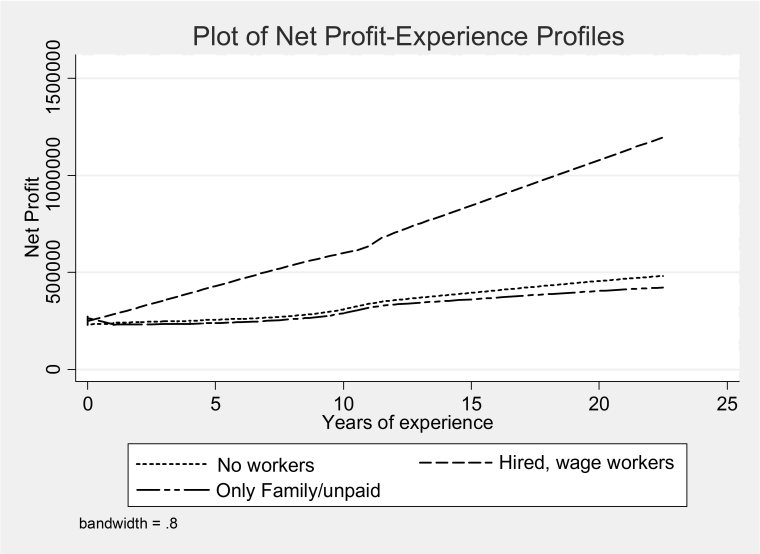

Figure 1: Plot of net profit against experience in three enterprise types

Note: The Figure records net profits of enterprises that startup in the three employment categories ((i) no employees, (ii) only family/unpaid workers, (iii) having waged employees) against years of experience of the individual entrepreneur running the enterprise.

A.2 Tables

Table 1: Summary Statistics on IFLS Rounds

Survey Round	Year	Individuals	Households	Enterprises	Communities
IFLS4	2007-08	44103 (50580)	13536	6186	313
IFLS3	2000	38433 (43649)	10435	5452	311
IFLS2	1997	22019 (33081)	7619	2625*	313
IFLS1	1993-94	22019 (33081)	7224	2439*	312
	Overall	66784	(unique)		

*In IFLS1 and IFLS1 household's only report on one, "primary" enterprise.

Table 2: Summary statistics on individual entrants

1998 Entrants

	N	Mean	SD	P25	Median	P75	P95	P99
Age	684	33.81	12.78	25	30	40	60	73
Marriage (married=1)	684	0.87	0.34	1	1	1	1	1
Gender (male=1)	684	0.71	0.46	0	1	1	1	1
Education (years)	684	5.58	6.04	0	0.5	12	15	19

1999 Entrants

	N	Mean	SD	P25	Median	P75	P95	P99
Age	1355	27.67	10.72	20	25	33	49	59
Marriage (married=1)	1355	0.87	0.33	1	1	1	1	1
Gender (male=1)	1355	0.68	0.46	0	1	1	1	1
Education (years)	1355	7.14	6.15	0	9	12	19	19

Table 3: Summary statistics on community-level sources of variation

Year 1998 Variable	N	Mean	SD	P25	Median	P75	P95	P99
Avg. change formal 1998	573	-0.35	1.03	-1.00	-0.75	0.00	2.00	3.00
Growth employment 1998	631	1.09	2.23	-0.25	0.33	1.75	5.50	9.00
Comm unemployment rate 1998	625	0.13	0.18	0.00	0.00	0.22	0.50	0.73

Year 1999 Variable	N	Mean	SD	P25	Median	P75	P95	P99
Avg. change formal 1999	526	0.03	0.19	0.00	0.00	0.00	0.50	1.00
Growth employment 1999	829	0.12	0.23	0.00	0.00	0.17	0.50	1.00
Comm unemployment rate 1999	1245	0.29	0.33	0.00	0.17	0.50	1.00	1.00

Table 4: Panel wage regression

VARIABLES	(1)
Age	-729.6043
	(970.2945)
Age^2	219.4695***
	(11.0449)
Education (years)	17,743.6683***
	(328.4670)
Marriage (married=1)	233,807.5485***
	(5,274.8745)
1989 year dummy	-15,202.2139**
	(6,129.3431)
1990 year dummy	-29,255.7173***
	(6,002.8286)
1991 year dummy	-43,050.8011***
	(5,913.6615)
1992 year dummy	-7,592.2594
	(5,578.9516)
1993 year dummy	308,289.2601***
	(12,759.3060)
1994 year dummy	471,856.2062***
	(5,852.7692)
1995 year dummy	423,124.5305***
	(5,578.6388)
1996 year dummy	365,138.9674***
	(5,215.6677)
1997 year dummy	334,528.8544***
	(5,050.2212)
1998 year dummy	151,353.3406***
	(4,807.7177)
1999 year dummy	142,859.3440***
	(4,393.3809)
2000 year dummy	156,652.8447***
	(4,177.0906)
Constant	-152094.1954***
	(20,436.4606)
Observations	75843
R-squared	0.490
R-squared adjusted	0.3515
F	3583.5

Note: Standard errors in parentheses

*** p <0.01, ** p <0.05, * p <0.1

Note: Observations restricted to 4.6 Rph. < wage < 1872075 Rph.

Table 5: Change in propensity to stay in self-employment after Asian Financial Crisis, 1998 entrants

Using 1998 New Entrants Into Self-Employment

Using Year-by-Year Choice Estimates

Year	Obs.	Mean	Std. Dev.	t-test	
1999	477	0.37	0.27	29.75	***
2000	477	0.33	0.28	26.19	***
2001	477	0.22	0.21	23.18	***
2002	477	0.22	0.23	20.76	***
2003	477	0.22	0.23	20.71	***
2004	477	0.24	0.23	22.89	***
2005	477	0.24	0.22	23.98	***
2006	477	0.24	0.22	23.23	***
2007	477	0.26	0.25	22.44	***
2008	477	0.32	0.25	27.67	***

Using Pooled Choice Estimates

Year	Obs.	Mean	Std. Dev.	t-test	
1999	477	0.34	0.26	28.64	***
2000	477	0.36	0.26	30.07	***

Using 1999, 2000 Choice Estimates

Year	Obs.	Mean	Std. Dev.	t-test	
1999	477	0.37	0.27	29.90	***
2000	477	0.33	0.28	26.22	***

Using Pooled Estimates Pooled Across 1999-2008

Year	Obs.	Mean	Std. Dev.	t-test	
1999-2008	954	0.35	0.26	154.16	***

Using 1999-2000 Pooled Across 1999-2000

Year	Obs.	Mean	Std. Dev.	t-test	
1999-2000	954	0.35	0.27	39.58	***

Note: Reports on the mean change in propensity to remain in self-employment for those who enter self-employment during 1998, with various methods of measuring propensity to persist and couterfactual. t-tests reported testing for differences in means of distributions.
*** p <0.01, ** p <0.05, * p <0.1

Table 6: Change in propensity to stay in self-employment after Asian Financial Crisis, 1999 entrants
Using 1999 New Entrants Into Self-Employment

Using Year-by-Year Choice Estimates

Year	**Obs.**	**Mean**	**Std. Dev.**	**t-test**	
2000	514	0.90	0.11	191.95	***
2001	514	0.42	0.15	63.44	***
2002	514	0.44	0.17	57.48	***
2003	514	0.45	0.16	62.64	***
2004	514	0.48	0.17	64.91	***
2005	514	0.52	0.17	68.29	***
2006	514	0.57	0.18	73.41	***
2007	514	0.66	0.18	83.30	***
2008	514	0.75	0.16	108.56	***

Using Pooled Choice Estimates

Year	**Obs.**	**Mean**	**Std. Dev.**	**t-test**	
2000	514	0.89	0.09	217.80	***

Using Pooled Estimates Pooled Across 2000-2008

Year	**Obs.**	**Mean**	**Std. Dev.**	**t-test**	
2000-2008	514	0.89	0.09	307.73	***

Note: Reports on the mean change in propensity to remain in self-employment for those who enter self-employment during 1999, with various methods of measuring propensity to persist and couterfactual. t-tests reported testing for differences in means of distributions.
*** p <0.01, ** p <0.05, * p <0.1

Table 7: Returns to experience, individuals entering self-employment during 1998

	Coefficient	Bootstrap Std. Err.	z	P > | z |
Exp. 1 curr. ent.	122053.6**	60520.25	2.02	0.04
Exp. 1 curr. ent.^{2}	-29698.09*	16522.57	-1.80	0.07
Exp. 2 curr. ent.	111070.5	82077.02	1.35	0.18
Exp. 2 curr. ent.^{2}	-19199.98	24561.68	-0.78	0.43
Exp. 3 curr. ent.	351829.4*	190108.40	1.85	0.06
Exp. 3 curr. ent.^{2}	-66646.76	52523.57	-1.27	0.20
Age	14752.11	11359.57	1.30	0.19
Age^2	-176.0679	134.49	-1.31	0.19
Gender	319052.6***	53910.98	5.92	0.00
Education (years)	19810.85***	3808.41	5.20	0.00
Marriage (married=1)	-1841.24	75149.90	-0.02	0.98
Selection correction, 1998 entry	-137276.9	116344.20	-1.18	0.24
Constant	-17804.76	220855.60	-0.08	0.94
Observations	1220			
Bootstrap replications	50			
Wald Chi-sq.	110.86			
Prob >Chi-sq.	0.00			
R-squared	0.0564			

*** p <0.01, ** p <0.05, * p <0.1

Note: Estimation by OLS with bootstraped standard errors due to projected regressor (selection correction).

Table 8: Returns to Experience, Individuals Entering Self-Employment During 1999

	Coefficient	Std. Err.	t	P >|t |
Exp. 1 curr. ent.	-158.00***	55.99	-2.82	0.005
Exp. 1 curr. ent. ^{2}	31.31**	14.55	2.15	0.032
Exp. 2 curr. ent.	316.40**	158.59	2	0.046
Exp. 2 curr. ent. ^{2}	-198.14**	89.32	-2.22	0.027
Exp. 3 curr. ent.	1623.77***	532.92	3.05	0.002
Exp. 3 curr. ent. ^{2}	-651.30**	299.85	-2.17	0.03
Age	5.76	39.35	0.15	0.884
Age ^{2}	-0.17	0.36	-0.47	0.64
Gender	145.40**	67.60	2.15	0.032
Education (years)	26.05**	11.24	2.32	0.021
Marriage (married=1)	-24.97	56.97	-0.44	0.661
Selection correction, 1999 entry	-1213.65	1974.65	-0.61	0.539
Constant	431.39	1113.93	0.39	0.699
Observations	599			
F	9.13			
Prob >Chi-sq.	0.00			
R-squared	0.1575			
Adj R-squared	0.1402			

*** p <0.01, ** p <0.05, * p <0.1

Note: Estimation by OLS with bootstraped standard errors due to projected regressor (selection correction).

Table 9: Summary statistics on firms, 2008, firms with no employees

Enterprises with no employees

	N	Mean	SD	P25	P50	P75	P95
Bus owned by household	2711	1.0	0.1	1	1	1	1
Pct owned by household	41	38.2	22.1	25.0	50.0	50.0	50.0
Bus. operated out. home	2711	0.8	0.4	1	1	1	1
Applied for permit	2711	0.0	0.2	0	0	0	0
Permit issued	100	1.0	0.0	1	1	1	1
Cost obtain permit	100	13014.2	35310.1	2.7	11.9	78.4	1081000.0
Unpaid labor startup	2711	0.2	0.5	0.0	0.0	0.0	1.0
Wage labor startup	2711	0.1	1.6	0.0	0.0	0.0	0.0
Total labor startup	2711	1.3	1.7	1.0	1.0	1.0	2.0
Startup capital	2251	409.9	2816.6	10.8	54.1	216.3	1406.0
Current unpaid labor	2711	0.0	0.0	0.0	0.0	0.0	0.0
Current wage labor	2711	0.0	0.0	0.0	0.0	0.0	0.0
Current total labor	2711	0.9	0.2	1.0	1.0	1.0	1.0
Current land assets	2711	324.2	3720.3	0.0	0.0	0.0	216.3
Current building assets	2711	182.6	1329.4	0.0	0.0	0.0	432.6
Current 4-wheel vehicle	2711	122.4	885.3	0.0	0.0	0.0	324.5
Current other vehicles	2711	101.2	301.4	0.0	0.0	0.0	757.1
Curr. other non-farm eq.	2711	84.3	414.0	0.0	8.7	43.3	324.5
Current total capital	2711	814.7	4467.1	5.4	37.9	346.1	2379.4
Unpaid labor shutdown	151	0.6	0.9	0.0	0.0	1.0	2.0
Wage labor shutdown	151	0.3	1.1	0.0	0.0	0.0	2.0
Total labor shutdown	151	2.0	1.3	1.0	2.0	2.0	4.0
Net profit	2637	679.2	1958.3	146.0	389.4	778.7	1946.8
Total revenue	48	448.5	515.5	108.2	324.5	648.9	1081.6
Total expense	43	251.4	266.8	108.2	108.2	324.5	648.9
Ent. products consumed	2660	92.9	254.9	0.0	13.0	75.2	389.4
Ent returns used by HH	2649	448.6	681.3	86.5	259.6	584.0	1349.8
Ent returns left over	2643	119.9	727.4	0.0	0.0	54.1	540.8
Total procure. of goods	643	188.0	556.0	5.4	21.6	86.5	1081.6
Total sales	88	410.1	711.2	13.5	64.9	405.6	2163.1
Total shared profit	54	304.8	394.2	2.7	64.9	584.0	1092.4
Unit returns to capital (%)	2259	983.6	45179.7	0.9	4.5	20.6	133.3
Unit returns to labor (USD)	2496	693.8	2005.1	155.7	389.4	778.7	1952.2
Net ch. labor since start	2711	-0.2	1.7	0.0	0.0	0.0	0.0
Net ch. capital since start	2251	450.1	5100.0	-21.6	0.0	135.2	1676.4

Note: Monetary values converted to 2008 US dollars. Dummy variables have decimal values removed.
Note: The three enterprise categories are mutually exclusive; in 2008 there are 6186 firms reported by IFLS households.

Table 10: Summary statistics on firms, 2008, firms with waged employees

Enterprises with waged employees

	N	Mean	SD	P25	P50	P75	P95
Bus owned by household	1149	0.9	0.2	1	1	1	1
Pct owned by household	72	44.6	17.8	33.0	50.0	50.0	75.0
Bus. operated out. home	1149	0.8	0.4	1	1	1	1
Applied for permit	1149	0.3	0.5	0	0	1	1
Permit issued	340	1.0	0.2	1	1	1	1
Cost obtain permit	340	8763.2	29239.7	10.8	54.1	216.3	1081000.0
Unpaid labor startup	1149	0.7	1.2	0.0	0.0	1.0	2.0
Wage labor startup	1149	2.2	6.3	0.0	1.0	2.0	6.0
Total labor startup	1149	3.8	6.4	2.0	3.0	4.0	9.0
Startup capital	1024	2021.4	7684.3	54.1	324.5	1622.3	6489.3
Current unpaid labor	1149	0.6	1.2	0.0	0.0	1.0	2.0
Current wage labor	1149	3.5	8.9	1.0	2.0	3.0	10.0
Current total labor	1149	5.1	8.9	3.0	3.0	5.0	12.0
Current land assets	1149	1519.9	8200.2	0.0	0.0	0.0	5407.7
Current building assets	1149	1534.9	7019.0	0.0	0.0	216.3	6489.3
Current 4-wheel vehicle	1149	1556.1	6032.4	0.0	0.0	0.0	8652.4
Current other vehicles	1149	277.3	768.7	0.0	0.0	32.5	1406.0
Curr. other non-farm eq.	1149	898.2	4066.3	10.8	108.2	540.8	3244.7
Current total capital	1149	5786.5	15666.2	119.0	1081.6	4326.2	26005.8
Unpaid labor shutdown							
Wage labor shutdown							
Total labor shutdown							
Net profit	1108	2749.0	6432.8	519.1	1297.9	2595.7	10123.3
Total revenue	30	1087.5	969.5	648.9	1081.6	1081.6	3893.6
Total expense	31	646.6	457.0	216.3	648.9	1081.6	1297.9
Ent. products consumed	1133	223.4	692.7	0.0	32.5	173.1	986.4
Ent returns used by HH	1129	1194.8	1838.5	259.6	648.9	1297.9	3893.6
Ent returns left over	1110	782.9	3263.5	0.0	64.9	540.8	2595.7
Total procure. of goods	457	1696.5	8365.4	27.0	108.2	540.8	6489.3
Total sales	72	4579.2	14121.5	39.2	200.1	2974.3	27038.7
Total shared profit	40	1442.8	4516.8	41.1	384.0	1081.6	3839.5
Unit returns to capital (%)	1037	13.2	100.7	0.3	1.0	4.5	60.0
Unit returns to labor (USD)	1108	614.3	1121.6	144.2	324.5	648.9	1946.8
Net ch. labor since start	1149	1.3	7.9	0.0	0.0	2.0	6.0
Net ch. capital since start	1024	3848.7	14910.6	-45.4	216.3	2109.0	22496.2

Note: Monetary values converted to 2008 US dollars. Dummy variables have decimal values removed.
Note: The three enterprise categories are mutually exclusive; in 2008 there are 6186 firms reported by IFLS households.