The labor market, the decision to become an entrepreneur, and the firm size distribution

Markus Poschke*
McGill University and CIREQ

August 8, 2012

Abstract

Why do some people become entrepreneurs, how do institutions affect this choice, and how does this affect the firm size distribution and aggregate productivity? This paper addresses this question using a matching model with occupational choice and heterogeneity in both ability as a worker and ex ante unknown productivity of firm start-ups. This rich setting allows to address effects of heterogeneity and diverse types of institutions, like labor market institutions, entry restrictions, taxes, which can possibly differ by firm size and thereby allow addressing informality. Importantly, the model allows for a comparatively flexible lower tail of the firm size distribution and can explain the existence and persistence of small, low-productivity firms with low profits: their owners have low outside options in the labor market. Key effects from a preliminary analysis are the following: labor market conditions affect incentives to start firms differently for workers and the unemployed, with repercussions on aggregate productivity; and they affect the expected value of firm creation due to the possibility of failure. Labor market frictions can have a new effect here: they shape prospective entrepreneurs' value of failure. Given that failure of new projects is common, they can strongly affect not only entry rates, but also the type of firms that enter.

Keywords. Occupational choice, firm entry, self-employment, matching models, selection, firm size distribution.

JEL Classification. E24, J23, J62, J64.

*markus.poschke@mcgill.ca

1 Introduction

Why do some people become entrepreneurs (and others don't)? Why do so many firms fail? Why are firms so heterogeneous? What determines which type of potential entrepreneurs actually start firms? With these questions in mind, this paper explores the importance of heterogeneity in ability, of institutions and of labor market conditions for the choice of becoming an entrepreneur, for the firm size distribution, and for resulting macroeconomic implications. Herein, the focus is different from most of the existing literature (see e.g. Hopenhayn and Rogerson 1993, Ljungqvist 2002, Bertola 2004), which has mainly analyzed the effect of labor market conditions and regulation on firms' employment policies and the value of entry. Instead, I focus on their impact on a person's own labor market prospects and through this on the choice of becoming (or not) an entrepreneur.

In this perspective, labor markets matter for the entrepreneurial choice decision in two ways. The first is obvious: being on the labor market as a worker is the alternative to becoming an entrepreneur. The second is more subtle: In a world where a large fraction of firms fail,[1] prospective entrepreneurs need to make their choice with the possibility of failure in mind. And insofar as this may mean looking for employment again, labor markets matter.

The importance of this effect will differ significantly across the population. For instance, a highly trained employee in a formal sector job may make a good entrepreneur – but also has a lot to lose from failure. A low-skilled worker employed in the informal sector has less to gain from entrepreneurship, but also puts less at stake. In some cases, entrepreneurship may even provide more stable revenues than on-and-off wage work. As a consequence, the effects of institutions and policies can have important interactions with heterogeneity in the population. This needs to be taken into account in the analysis of both policies aimed at encouraging prospective high-productivity entrepreneurs and policies targeted at the subsistence self-employed.

To capture all these margins, I introduce an occupational choice decision into a labor market search and matching model where ability as a worker and the productivity of startups are heterogeneous, and the latter is unknown ex ante. Although there is some literature on the choice among occupations by workers, starting with Miller (1984), to the best of my knowledge the only previous treatment of the choice between working and entrepreneurship in the context of a matching model, i.e. with a focus on labor market determinants and outcomes, is Fonseca, Lopez-Garcia and Pissarides (2001), who focus on the impact of

[1] In the U.S., the probability of surviving the first four years is only 63%. It is similar in Chile and Argentina, and even substantially lower in Mexico, all according to Bartelsman, Haltiwanger and Scarpetta (2009). For more data on firm turnover and the dynamics of firms more generally, see e.g. Caves (1998), Foster, Haltiwanger and Krizan (2001), Tybout (2000) and references therein.

costly entry regulation on employment. I take an approach that is more general in several dimensions, and allows for richer results.

In the world modeled here, people differ in both productive ability (output as workers) and the productivity of firms they start. Whereas productive ability is known, the productivity of entrepreneurial projects can only be found out by implementing them, i.e. by becoming an entrepreneur. Average productivity of firms started by a person may however be correlated with his productive ability. Workers (and in an extension, also the unemployed) then decide whether to start firms. As a result, some but not all people try their luck as entrepreneurs, and some succeed, with the entry decision taken on the base of people's expectation of their potential firm's productivity, conditional on their observed productive ability.[2]

The main results of this paper are the following. Firstly, frictions in the labor market matter for occupational choice. For instance, facing longer expected unemployment duration in the case of failure with the firm discourages workers from entrepreneurship. Secondly, when the unemployed can also start firms, a lower value of unemployment (for instance, because finding a job takes a long time) encourages them to start firms. Indeed, in the data, the cross-country correlation between the level of firing costs (usually associated with longer unemployment duration, therefore lower value of unemployment) and the fraction of entrepreneurs starting a firm "out of necessity" is positive, giving some empirical support to this result (see Section 7.2). Moreover, if firms started by the unemployed are on average less productive as suggested by the data,[3] then this change in the composition of entrepreneurs reduces aggregate productivity. Thirdly, the correlation between productive ability and expected productivity as an entrepreneur significantly shapes the firm size distribution. Empirically observed distributions suggest a correlation that is positive but not too close to one. [More quantitative work could bound this more precisely – in progress.] Finally, the model can explain the existence and persistence of small, low-productivity firms with low profits. They persist because their owners' outside option in the labor market is even less attractive.[4] This is in contrast to most existing models with a firm size distribution (see e.g. Hopenhayn 1992, Luttmer 2007). They usually feature a unique exit threshold that

[2] A technical complication that needs to be dealt with are strategic considerations in hiring. As in the original contribution by Stole and Zwiebel (1996), firm-level decreasing returns to scale here inject a strategic element into hiring decisions by multiworker firms. Felbermayr and Prat (forthcoming) introduce firm heterogeneity into such a setting, and Cahuc, Marque and Wasmer (2004) allow for heterogeneous though not perfectly substitutable (as here) labor. Wage determination here is hence close to these papers, although they allow for less dimensions of heterogeneity, and do not analyze occupational choice and its impact on the firm size distribution. Cagetti and De Nardi (2006) do look at entrepreneurship and model people as having distinct entrepreneurial and working ability (both observable), but focus on liquidity constraints, not on labor market issues.

[3] For an extensive survey on the determinants of occupational choice, see Santarelli and Vivarelli (2006).

[4] For more evidence and a theoretical investigation of this pattern, see Poschke (2008).

bounds firm size and profits from below; here this threshold depends on each entrepreneur's outside option, which may be very low, allowing for the existence of small firms and a more flexible lower tail of the firm size distribution.

The paper is organized as follows. The next section sets out the economy. Section 3 analyzes bargaining and wage determination, and Section 4 occupational choice and the determination of the firm size distribution. Equilibrium conditions are summarized in Section 5. A closer look at the distribution of abilities and their impact on the firm size distribution is provided in Section 6. Finally, in Section 7, the model is extended to firm creation by the unemployed in order to analyze the impact of labor market conditions on occupational choice. The section closes with a brief glance at the data.

2 The economy

Time is discrete. The economy consists of a continuum of risk-neutral individuals of measure 1. At the beginning of every period, a fraction λ of these individuals retire, and a fraction λ newly enters the economy, so that the size of the work force remains constant. When an entrepreneur retires, the firm's employees become unemployed. People derive utility from consumption, and can earn income either as workers or by running their own firm. In periods of unemployment, they can attain a consumption level of b by non-marketable home-production activities. They discount future utility with a discount rate $\beta \in (0, 1)$, implying an interest rate of $\tilde{r} = (1 - \beta)/\beta$.

Ability. Individuals differ in their productive ability a, and firms differ in productivity α. Both attributes are constant over the life of the individual or the firm. A firm's productivity depends on the quality of its founder's idea and management. Suppose that more productive individuals have a tendency to have better ideas, so the correlation ρ between a group of individuals' productive ability and the productivity of the firms they might run is non-negative. Concretely, suppose that people's productive ability a and the productivity of firms they may operate are drawn from a bivariate normal with correlation $\rho \geq 0$. For notation, let $a \sim N(\mu_a, \sigma_a^2)$ and $\alpha \sim N(\mu_\alpha, \sigma_\alpha^2)$. Denote these marginal pdfs by $f_a(a)$ and $f_\alpha(\alpha)$, respectively, and the density of α conditional on a by $f_{\alpha|a}(\alpha|a)$. The expectation of α conditional on a then is linear in the realization of a, and the slope depends on ρ and the two variances. Finally, as the variance of income and wealth is significantly higher for entrepreneurs than for the population as a whole (Cagetti and De Nardi 2006), assume that $\sigma_\alpha^2 > \sigma_a^2$.

Firms produce output with the production function

$$y(\alpha, n) = \alpha n^{\gamma}, \quad 0 < \gamma < 1 \tag{1}$$

combining a manager/owner who is essential to operate a technology of productivity α, and workers n, with decreasing returns to scale in labor so that firm size is finite for all α. All firms produce a homogeneous good, which is used as the numéraire. Productive ability (or "ability" for short, to distinguish it from productivity, which is a firm-level concept) here means that a worker with ability a is equivalent to a workers of productivity 1, i.e. workers are perfect substitutes. A firm's total employment then is the integral over the a's of its employees.

A person's productive ability and quality as a manager may be correlated ($\rho > 0$), but are not the same. In fact, the literature on entrepreneurship (see in particular Lazear 2004) stresses that entrepreneurs need to be "Jacks-of-all-Trades" and have broad talents, as opposed to more specialized skills of production workers. However, previously employed entrepreneurs tend to be more successful than those coming from unemployment, and managers more than others, so it seems reasonable to suppose that there is a positive correlation between α and a.

Now suppose that when people enter the labor market, their productive ability a is known and observable (e.g. from their education), so it can be taken into account in wage setting. In contrast, the productivity α of a potential project is necessarily unknown (including to the entrepreneur), and can only be found out by starting the project. So a potential entrepreneur, knowing ρ and his a, only knows the distribution of α conditional on his a that he faces.

Firms and vacancies. All individuals can create a firm by making a sunk investment of k_f. When the project starts, its productivity α is revealed. Depending on its level, the entrepreneur may decide to pursue the project and to post vacancies to recruit employees, or not to pursue the concern and to become unemployed instead, starting to look for a job in the following period. Successful entrepreneurs operate their firm until retirement.

Suppose for now that only workers can start firms (but have to leave their job for that). Evidence suggests that workers are indeed better at starting firms, so the assumption seems not too restrictive. In an extension (see Section 7.1), the situation where both workers and the unemployed can start firms will be analyzed. Depending on the labor market situation, the incentives of the two groups can vary. For instance, when finding a job is difficult, this increases the incentive of the unemployed to start a firm, while decreasing that of workers to leave their job for a firm. If workers are better at running firms than the unemployed, labor market conditions will then affect the productivity of firms in operation by affecting the composition of entrepreneurs.

The labor market. Matching between workers and jobs occurs on a labor market that is characterized by search and matching frictions. Unemployed workers search for vacant jobs. Firms can post vacancies to fill vacant jobs, at a per-period vacancy posting cost of k_v. Suppose, in line with most of the literature, that matching between unemployed workers u and vacancies v can be described by a constant-returns Cobb-Douglas matching function. Then the number of matches made in a period, M, is given by $M = Au^{\mu}v^{1-\mu}$, where A is the productivity of the matching technology, or matching efficiency. Defining labor market tightness θ as the vacancy-unemployment ratio v/u, the probability of filling a vacancy, M/v, is then given by $q \equiv A\theta^{-\mu}$. For unemployed people, the probability of finding a job equals $M/u = \theta q$. A tighter labor market means that there are more vacancies per unemployed person, and it becomes easier to find a job (for the unemployed), and more difficult to fill a vacancy (for firms). Because of the frictions inherent in the matching setup, not all unemployed people are matched with a new job. Some remain unemployed and will again search for a job in the subsequent period.

Workers are heterogeneous. However, the focus here is not on random differences across firms in the productivity of the workers they employ just because of differences in whom they meet on the labor market. Therefore, assume the following: Firms post vacancies for, and meet if successful, "effective workers" of ability $\bar{a}$, where $\bar{a}$ is the average ability of workers in the pool of the unemployed. Moreover, the job finding and vacancy filling probabilities are independent of a worker's ability or a firm's productivity. These assumptions clearly stretch the search and matching framework a long way. However, they have the great advantage of allowing to maintain a role for worker heterogeneity in the occupational choice problem, and only there, eliminating its effect on the firm's employment problem. (The assumption is also a way of breaking down the firm's employment of a continuous, perfectly substitutable workforce into "natural" units.)

When a firm and a worker meet, they engage in Nash bargaining over the surplus created by a job, with the worker's bargaining power given by η and the firm's by $1 - \eta$. It is shown below that this can be satisfied for all workers, not only for the effective worker, and that wages increase in a. As shown by Stole and Zwiebel (1996), multiworker firms have an incentive to overhire. The effective worker assumption also allows to deal with this complication despite the heterogeneity of workers.

From here it is clear that this economy extends the standard search and matching setup as in Pissarides (2000) by heterogeneous multiworker firms, heterogeneous workers, and agents' occupational choice decision. While some of these elements have been combined in the literature and there now is a substantial literature on search and matching with multiworker firms, the occupational choice decision has not received much attention. To the best of my knowledge, it has only been employed in Fonseca et al. (2001) in an analysis of the impact of firm entry barriers on employment, however, with heterogeneity only in one dimension (known entrepreneurial ability), and with the restriction that only unemployed people could create firms.

The occupational choice problem has the following basic structure. Workers find it optimal to start a firm if this yields higher value than their current employment. The latter depends on the worker's productive ability. This is also true for the former, because more able workers also tend to run more productive firms. As a result, it may be that not all workers start firms; which ones do so will crucially depend on the correlation between α and a. Once a firm has been started, the owner will only pursue the concern if this gives a higher value than giving up and looking for a job again. This implies a minimum productivity needed for the firm to continue. As the outside option depends on a worker's ability, so does the threshold. This threshold determines a venture's success probability and thereby also affects the startup choice.

Before pursuing the occupational choice problem in detail, it is necessary to find wages and the firm value. So the next section, after a brief statement of the (largely standard) value of the different states, analyzes the firm's employment decision. The following one then proceeds to the occupational choice decision.

3 Matching, employment, and wages

Upon meeting, the worker and the firm bargain over the division of the match surplus. For a worker with productive ability a, the value of being employed W consists in the wage and either the value of employment in the next period, if the firm continues, or the value of unemployment if the firm closes down because the owner retires (which occurs with probability λ). Here and in the following, all values are measured at the end of the period, and $1/(1+r) = (1-\lambda)/(1+\tilde{r})$ is the discount rate that incorporates the agent's own retirement probability λ.

$$W(a) = \frac{1}{1+r}\left(w(a) + (1-\lambda)W(a) + \lambda U(a)\right). \tag{2}$$

This expression omits the option value of starting a firm gained by having employment. (Remember that only workers can start firms.) Suppose for now that both the firm and the worker (also later, as an entrepreneur making his continuation decision) ignore this.

[Assumption to be relaxed.] If this option had positive value, and if the firm could take this into account in bargaining, it could use this to bid down the wage. However, under the assumption that the firm hires effective workers and cannot distinguish who is going to start a firm, this is not possible. In any case, in a more general case where both employed and unemployed people can start firms (considered in an extension below), this effect would again be attenuated. On the workers' side, the main effect of the assumption that they ignore the option value is to make them less discriminating in the choice of projects they decide to pursue. This is because they do not take into account that the outside option is not simply the value of unemployment but also incorporates the option of starting a new firm after finding employment. Qualitatively, the assumption does not change any of the results derived in the following. Quantitatively, it mainly implies less strict selection.

The value to a worker of being unemployed, $U(a)$, then consists in immediate utility b and in the discounted value of either unemployment $U(a)$ or employment $W(a)$ next period, where θq is the probability of finding a job. Through $W(a)$, the value of unemployment also depends on a person's productive ability a.

$$U(a) = \frac{1}{1+r} \left(b + \theta q W(a) + (1-\theta q) U(a) \right). \tag{3}$$

So the worker's gain from finding a job is

$$W(a) - U(a) = \frac{w(a) - rU}{r+\lambda}. \tag{4}$$

A firm maximizes its value F by choosing the number of vacancies v to post:

$$F = \max_v \frac{1}{1+r} \left(y(\alpha, n) - \int w(a) n(a) \mathrm{d}a - k_v v + F(n') \right), \tag{5}$$

where the expression for the wage bill integrates over all types of workers, and $F(n')$ is firm value with the new employment level n'. Employment evolves according to

$$n' = (1-\lambda)n + qv. \tag{6}$$

The marginal value of an additional worker in this situation consists not only of the usual terms of marginal product, wage, and marginal value next period, but also of his impact on the total wage bill. The reason is that with decreasing returns, the marginal worker has lower marginal product than previous workers. Suppose that the firm bargains individually with each worker; then it can treat each of them as the marginal worker. Stole and Zwiebel (1996) show that their wage then is a weighted average of marginal products. Hiring an additional worker reduces this weighted average, so new hires have an effect on other workers' wages which the firm can take into account, giving rise to an overhiring effect. Although this is not in focus here, it does occur. The situation considered here comes closest to Cahuc et al. (2004) who provide a general equilibrium treatment of the problem allowing for several inputs. They do not, however, allow for heterogeneous but perfectly substitutable workers, so our solution will differ from theirs in the use of the effective worker assumption.

Since a firm's productivity is constant over time and aggregates are constant, too, it is optimal to choose a constant employment level $n(\alpha)$. For this, the firm has to replace $\lambda n(\alpha)$ retired workers every period by posting q times as many vacancies. To reach employment n, the firm needs to employ $\bar{n}$ effective workers with productive ability $\bar{a}$ each, so $n = \bar{n}\bar{a}$. Let these workers' wage be $\bar{w}(\bar{n})$: it depends on the firm's total employment of effective workers due to the presence of decreasing returns. The wage bill then is $\bar{w}\bar{n}$, and the marginal wage cost of an additional worker is $\bar{w}+\bar{n}\bar{w}'(\bar{n})$; the worker's own wage plus the effect of hiring him on all other wages. Using this and constancy over time, the marginal value of an additional effective worker can be solved for as

$$F'(\bar{n}) = \frac{1}{r}\left(\bar{a}y'(\alpha, n) - \bar{w} - \bar{n}\bar{w}'(\bar{n})\right). \tag{7}$$

Note that the marginal product is multiplied by $\bar{a}$ because by hiring an additional effective worker, n is changed by $\bar{a}$. An expression for the firm's labor demand can then be obtained from the first order condition for vacancies as

$$\bar{a}y'(n) = \bar{w} + \bar{n}\bar{w}'(\bar{n}) + k_v/q. \tag{8}$$

The firm equates the marginal product of an effective worker to his effect on the wage bill, plus the expected recruitment cost.

In the usual way, Nash bargaining gives rise to the condition that the wage is set such that the surplus is split according to the bargaining weights:

$$(1-\eta)\frac{\bar{w} - rU(\bar{a})}{r+\lambda} = \eta\frac{F'(\bar{n})}{r}. \tag{9}$$

Substituting in equation (7), the wage then is a weighted average of the worker's outside option and his effect on the firm's output and wage bill.

$$\bar{w} = (1-\eta)\frac{r}{r+\eta\lambda}rU + \eta\frac{r+\lambda}{r+\eta\lambda}(\bar{a}y'(n) - \bar{n}\bar{w}'(\bar{n})). \tag{10}$$

The solution to this linear first order differential equation is[5]

$$\bar{w} = \frac{\eta(r+\lambda)}{(1-\eta)r}\frac{k_v}{q(\theta)} + rU(\bar{a}). \tag{11}$$

This is the wage that solves the Nash bargaining problem if the firm always employs as indicated by its labor demand function given in equation (8).[6] It shows that the wage does not depend on a firm's productivity. This is because a multiworker firm that bargains with all its workers simultaneously and treats each as the marginal one can extract all surplus above the worker's outside option plus a share of the recruitment cost that would be needed to hire someone else, and these quantities are the same across firms.

While equation (11) gives the wage for an effective worker, the wage for workers with different a is a linear deviation around this. This follows from two first-order approximations valid for "small" workers, as is the case with the continuum here: The firm's marginal product of labor is linear in a (at $ay'(n)$, thus supposing no decreasing returns *within worker*), and the effect on the total wage bill is also linear (at $\int an(z)\partial w(z)/\partial n(a)\mathrm{d}z$.) Then the only other term in a in the Nash bargaining rule is the wage, so this also has to be linear in a. By substituting the firm's demand for effective workers into the sharing rule for general a, it follows that

$$w(a) = \frac{r(1-\eta)}{r+\eta\lambda}rU + \frac{(r+\lambda)\eta}{r+\eta\lambda}\frac{a}{\bar{a}}\left(\frac{k_v}{q} + \bar{w}\right). \tag{12}$$

How much the wage varies with a thus depends on recruiting cost and on the average wage (this comes from perfect substitutability), whereas the lower bound is related to the outside option.

4 Firm value, entry, and survival

Using the definition of the production function, the firm's labor demand for effective workers then is

$$\bar{n}(\alpha, \bar{a}, \theta) = \left[\frac{c_1(\gamma-1)+1}{\alpha\gamma\bar{a}^{\gamma}}\left(\frac{k_v}{q} + \bar{w}\right)\right]^{\frac{-1}{1-\gamma}}, \tag{13}$$

[5]See the appendix for the derivation. Cahuc et al. (2004) give detailed solutions for a similar case and for cases with multiple inputs.

[6]If a firm's employment is not constant equation (7) cannot be used, and the expression for the wage can differ. This can occur for instance because entrants can only grow slowly because recruitment within a period is limited as in Acemoglu and Hawkins (2006), or if there are non-linear adjustment costs, so entrants prefer to expand slowly.

where $c_1 = \eta(r+\lambda)/(r+\eta\lambda)$. With decreasing returns, labor demand is convex in productivity α, a property that carries through to profits and firm value

$$F(\alpha) = \frac{1}{r}\left(y(\bar{n}\bar{a}) - (\bar{w} + \lambda k_v/q)\bar{n}\right). \tag{14}$$

This implies that the owner of a startup with productivity α and own productive ability a only continues if $F(\alpha) \geq U(a)$, implying a reservation productivity

$$\alpha_R(a) = \{\alpha | F(\alpha) = U(a)\}. \tag{15}$$

Hence, a startup owned by someone with ability a continues if it draws a productivity $\alpha \geq \alpha_R(a)$. Because F is convex in α whereas U is linear in a, α_R is concave in a: the marginal value of α is increasing, so as the values rise, smaller increases in α are needed to compensate the effect of increases in a on U. (If the owner took into account that quitting allows trying again later on, α_R would be higher for given a, as discussed above.)

Before entry, the worker knows the conditional distribution of α for his level of productive ability. So he also knows the probability of failure $d(a)$, i.e. the probability that the draw of α will lie below the threshold $\alpha_R(a)$. As the latter depends on a, so does $d(a)$. Since the conditional variance of α at $(1-\rho^2)\sigma_\alpha^2$ does not depend on a, the failure probability is monotonically increasing in the difference $\alpha_R(a) - \mathsf{E}(\alpha|a)$. Because α_R is concave and the conditional expectation increases linearly, the difference is hump-shaped, implying that the failure probability peaks together with the difference.

Knowing the conditional distribution, a prospective entrepreneur also knows the distribution of possible payoffs in the case of success. The expected value of entry conditional on productive ability a hence is

$$\mathsf{E}[F(\alpha)|a] = -k_f + \int f(\alpha|a)\max(U(a), F(\alpha))\mathrm{d}\alpha \tag{16}$$

$$= -k_f + d(a)U(a) + (1-d(a))\mathsf{E}[F(\alpha)|a, \alpha \geq \alpha_R]. \tag{17}$$

It consists in the cost of entry, the value of unemployment in the case of failure, and the expected value of the firm in the case of success. As it is a weighted average of linear and convex objects, it is itself (weakly) convex.[7] It is also useful to know that is bounded below by the value of unemployment minus the entry cost.

[7]To see this, notice that $\max(U, F)$ is weakly convex, and that the integral averages over these weakly convex objects. The expected value cannot be linear everywhere unless parameters are such that the probability of failure is 1 for all a. Then however there would be no entry, hence no firms. But this also implies zero labor demand and no employment. But then, unemployment cannot be preferable to entry – this cannot be an equilibrium.

When considering entry, workers compare the value of employment to the expected value of the firm, and enter if $\mathsf{E}[F(\alpha)|a] \geq W(a)$. Note here that the possibility of failure, resulting in unemployment, just acts as another entry cost. Moreover, since the difference between W and U increases with the wage and therefore with a, this barrier matters more for high-wage workers. Despite (in expectation) being better managers, they have more to lose by giving up their job. Due to the convexity of the value of entry, this does not matter in the limit – there are levels of productive ability associated with high enough expected productivity that entry is worthwhile – but it may matter in some regions.

In the aggregate, two different outcomes of the occupational choice problem can arise, with the difference concerning low-a workers. In general, the linear $W(a)$ curve can cut the convex $\mathsf{E}(F|a)$ curve twice or miss it. In the last case (as in the case of tangency) a measure zero of people would choose to work, but then entrepreneurs would not find workers to employ, so being an entrepreneur would not be profitable. So this case cannot be an equilibrium. In the case of two crossings, let the values of a where $W(a) = \mathsf{E}(F|a)$ be a_L and a_H, with $a_L < a_H$. Both or only one of them can lie in the admissible, positive domain of a. If only one lies here, only high-a workers ($a \geq a_H$) start firms. If both occur here, workers with extreme values of a ($a \leq a_L$ or $a \geq a_H$) start firms. Why high-a workers start firms is obvious: the value of a firm increases convexly in α, the distribution of α conditional on a improves linearly with a, so expected firm value increases convexly, and faster than the value of employment. So there will always be levels of a high enough that workers start firms. For low a, the reason is the opposite: Although these workers only expect to run rather small firms, in a world of decreasing returns these are not much less profitable than their medium counterparts, while the wage as an employee declines faster with a in this range. Said the other way around, these workers have so low ability that even running a low-productivity firm is preferable to employment.

A graph is useful for understanding the decision. Figure 1 plots the relevant quantities. The two rows of plots show outcomes for two different levels of the entry cost.[8] To follow the occupational choice decision, start with the middle column. It shows the reservation productivity $\alpha_R(a)$ (the concave line) and expected productivity $\mathsf{E}(\alpha|a)$ (linear) as functions of a. The reservation productivity rises in a because the value of the worker's outside option (unemployment) increases in a. A startup's expected productivity also increases in a if $\rho > 0$. The probability of failure $d(a)$ (shown in the last column) follows from this as the probability of drawing a value of α below the reservation productivity, conditional on a.

[8] It is 0.2 in the first row, 10 in the second. The other parameters are $A = 0.1, b = 4, k_v = 1.5, \mu_a = 3, \mu_\alpha = 25, \sigma_a = 1, \sigma_\alpha = 4, \rho = 0.75, \eta = 0.5, \mu = 0.5, \gamma = 2/3, r = 5\%$ yearly, λ such that expected working life is 40 years. This is a rough, preliminary parametrization; obtaining a better fit to the firm size distribution is an interesting objective.

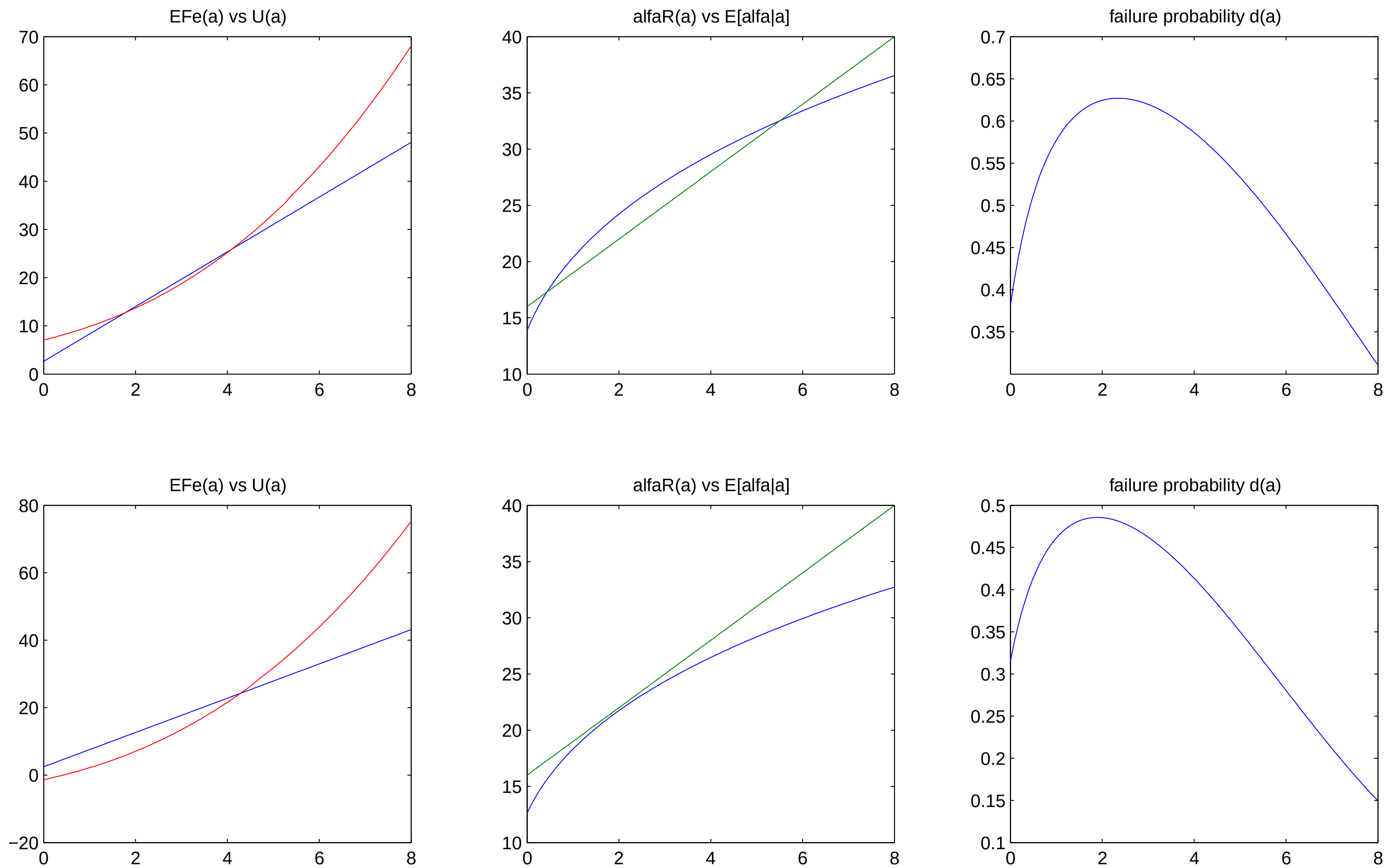

Figure 1: The occupational choice decision

Due to the non-linearity of $\alpha_R(a)$, the failure probability is non-monotonic in a. This is not general: depending on parameters, $d(a)$ can also be monotonic, and increasing or decreasing in a, as discussed below when analyzing the effect of ρ. These elements, combined with firm value in case of success, determine the expected value of entry as a function of a, which is plotted in the first column, together with the linear value of employment. This figure clearly shows for which a it is preferable to start a firm, and for which it is preferable to remain an entrepreneur. The intersections define a_L and a_H. The two rows illustrate the two possible cases: one where low-a workers do decide to start a firm, and one where they do not. High entry costs as in the lower row of plots, among other things, can make this option unattractive.

In this way, the model explains a feature of the firm size distribution not commonly well explained in existing models: despite low profits, small firms are a persistent feature of the firm size distribution because their owners have even lower outside options in the labor market. This effect is particularly strong in industries with low entry costs. The result is also consistent with a substantial fraction of firms making little profits, as documented by Moskowitz and Vissing-Jørgensen (2002) (and dubbed a possible "Private Equity Premium Puzzle"). While they reject quite a number of possible explanations of this pattern, a low outside option is a new candidate. A better quantitative characterization of the firm size and profit distribution in the model here [in progress] would help assessing how far it goes. Hintermaier and Steinberger (2005) propose a related but different explanation where low returns essentially constitute an investment by the entrepreneur in information acquisition about his own unknown entrepreneurial ability. This is similar to the element of experimentation present here (see below) but quite different from the effects of heterogeneous ability. They would not predict low-profit firms to persist, which does however occur. (In the model, low-profit firms started by high-a entrepreneurs would not continue, but those started by low-a people would, resulting in exit rates that decrease in α, but still some survival of low-α firms.

For later use it is useful to show that $d(a_H) < d(a_L)$ when both thresholds exist. At the threshold, $\mathsf{E}(F|a) = W(a)$. Substituting in from (16) and solving for d gives the failure probability at the threshold as

$$d = \frac{\widetilde{EF} - W - k_f}{\widetilde{EF} - U} = \frac{\widetilde{EF} - W - k_f}{\widetilde{EF} - W + W - U} \tag{18}$$

where $\widetilde{EF}$ is the expected payoff in the case of success, $\mathsf{E}[F(\alpha)|a, \alpha \geq \alpha_R]$. Since $W(a) - U(a) = (w(a) - b)/(r + \lambda + \theta q)$ increases in a, $d(a_H) < d(a_L)$. Having more to lose, high-a entrepreneurs require a higher success probability to start a firm. (Note that the condition only holds at the thresholds; d does not generally decline in a.)

The thresholds a_L and a_H define the law of motion of the distributions of firms, workers, and unemployed. As a person's a is constant and the thresholds sort individuals, people with a between the thresholds will always be either employed or unemployed, but will never start a firm. People with a below a_L or above a_H, meanwhile, will always try staring firms, but may also be unemployed due to failure with a startup, or working and waiting to start a firm in the next period.

Indicating the change in a variable by a dot above it, the distribution of entrepreneurs then evolves according to

$$\dot{e}(a) = (1-\lambda)n(a|a \leq a_L \vee a \geq a_H)(1-d(a)) - \lambda e(a). \tag{19}$$

Every period, surviving workers with a below a_L or above a_H attempt entry and succeed with probability $1-d(a)$, while a fraction λ of entrepreneurs retires. The distribution of workers evolves according to

$$\begin{aligned}\dot{n}(a) = {} & [\lambda f_a(a) + (1-\lambda)u(a)]\theta q - \\ & - n(a|a \leq a_L \vee a \geq a_H) - n(a|a_L < a < a_H)\left(\lambda + (1-\lambda)\lambda\right).\end{aligned} \tag{20}$$

New jobs are taken up by entrants into the labor force and surviving unemployed workers who find a job, while workers who start a firm, retire, or are laid off leave employment. Finally, the evolution of unemployment is given by

$$\begin{aligned}\dot{u}(a) = {} & \lambda f_a(a)(1-\theta q) + n(a|a_L < a < a_H)(1-\lambda)\lambda + \\ & + (1-\lambda)n(a|a \leq a_L \vee a \geq a_H)d(a) - u(a)(\lambda + (1-\lambda)\theta q).\end{aligned} \tag{21}$$

There are three ways of entering unemployment here: entering the labor force and not finding a job, being laid off, or failing with your firm. People exit unemployment via retirement or by finding a job. As a person's a is constant, these three equations can easily be solved for the distributions of productive ability of firm owners $e(a)$, employed workers $n(a)$, and the unemployed $u(a)$. The latter determines average productivity of the unemployed, $\bar{a}$, that firms face in their hiring decisions. The total measure of entrepreneurs e, employed workers n, and unemployed u can be obtained by enforcing the constraint that they have to sum to the size of the labor force, 1.

While the distribution of firms over their owners' ability a is not very useful to know in itself, it implies the firm productivity distribution, which allows obtaining aggregate labor demand and vacancy posting, and the firm size distribution. Denote the measure of firms with productivity α by $\nu(\alpha)$. For any fixed α it can be found by using the firm distribution over a and the conditional distribution of α and integrating out the a, so that

$$\nu(\alpha) = \int e(a) f_{\alpha|a}(\alpha|a, \alpha > \alpha_R(a)) \, \mathrm{d}a. \tag{22}$$

Here, $f_{\alpha|a}(\alpha|a, \alpha > \alpha_R(a))$ is the probability of α conditional on a if α exceeds the reservation productivity associated to this a, and zero otherwise. In this way, $\nu(\alpha)$ captures only firms whose owners find it optimal to continue.

Knowing this distribution, the model is closed in the labor market. Tightness, the ratio of vacancies to unemployment, can be obtained using firms' labor demand. Each continuing firm needs to hire $\lambda\bar{n}(\alpha, \bar{a}, \theta)$ workers to replace retiring workers. Each entering firms wants to hire $\bar{n}(\alpha, \bar{a}, \theta)$ workers to reach its optimum employment level. To achieves this in expectation, the number of vacancies firms post is $1/q$ times desired hiring, as indicated by the first order condition for vacancies. Total vacancies are then given as

$$v(\bar{a}, \theta) = \frac{\lambda + (1-\lambda)\lambda}{q} \int \nu(\alpha)\bar{n}(\alpha, \bar{a}, \theta) \, \mathrm{d}\alpha. \tag{23}$$

Remember that labor demand and occupational choice depend on the tightness θ and the average productivity of the unemployed $\bar{a}$ that firms take as given. These decisions at the firm level in turn determine θ and $\bar{a}$. For an equilibrium, the θ and $\bar{a}$ implied by the firms' decisions need to equal those underlying those decisions.

5 Equilibrium

To establish existence and uniqueness of equilibrium, it is convenient to reason in terms of θ and the wage of the effective worker, $\bar{w}$, instead of his productivity $\bar{a}$. As in standard search and matching models, the wage here is an increasing function of tightness for all a, whereas firms' labor demand, due to recruitment cost, decreases in tightness (see equations (11) and (13)). By making work more attractive relative to running a firm, tightness also reduces the number of firms, making sure that both firm-level and aggregate labor demand fall in tightness. As a consequence, the model has a unique equilibrium if b is not too high, i.e. as long as the productivity of non-market activities is low enough compared to market work, so that bargained wages are such that firms would indeed choose positive employment at that wage.

A stationary equilibrium is then characterized by quantities $\bar{a}, v, u, e, n$, prices $\bar{w}$ and $w(a)$, value functions $U(a), W(a), F(\alpha), \mathsf{E}[F(\alpha)|a]$, probabilities $d(a)$, a function $\alpha_R(a)$, thresholds

a_L and a_H, and distributions $e(a), n(a)$, and $u(a)$ such that

- the average wage $\bar{w}$ (equation (11)) solves the bargaining problem between the worker and the firm given in equation (9) for an effective worker, and wages $w(a)$ (equation (12)) solve it for all other a, taking θ and $\bar{a}$ as given;
- the firm posts vacancies and chooses employment optimally, taking θ and $\bar{a}$ as given;
- values $U(a), W(a)$, and $F(\alpha)$ are as given by equations (2), (3), and (14) for all θ and $\bar{a}$;
- firms that just entered optimally decide upon continuation, defining the reservation productivity $\alpha_R(a)$ as given in equation (15), taking θ and $\bar{a}$ as given;
- prospective entrants correctly perceive the failure probability $d(a)$ as $Pr(\alpha < \alpha_R(a))$, taking θ and $\bar{a}$ as given;
- they evaluate the expected value of entry as given in equation (16), taking θ and $\bar{a}$ as given;
- they optimally decide whether to enter or continue working, defining the thresholds a_L and a_H, taking θ and $\bar{a}$ as given;
- the distributions of firms, unemployed, and workers evolve according to equations (19) to (21);
- the sum of entrepreneurs, unemployed, and workers adds up to the labor force;
- the distribution of firms over α is as defined in equation (22);
- aggregate vacancies are posted as given in equation (23) for every θ and $\bar{a}$;
- and tightness θ and the average productivity of the unemployed implied by the equilibrium measure and distributions of firms and the unemployed equal tightness and the productivity of the effective worker $\bar{a}$ taken as given by firms and workers in their optimization problem.
- Finally, the equilibrium is stationary, i.e. all equilibrium objects are constant over time.

6 The allocation of talent

One important feature of the economy described here is the influence of ρ, the correlation between productive ability and the quality of the ideas generated, on occupational choice and welfare. Note in passing that ρ is a parameter of technology, not information, so this is not an issue of imperfections in some sense, but of productivity of the available technology.

Now consider some extreme cases, illustrated in Figure **??**.[9] For low ρ, illustrated here with $\rho = 0.5$ in the first row of Figure **??**, high- and low-a workers face very similar quality of projects they could generate. $\mathsf{E}(\alpha|a)$ is not very steep. Since high-a workers have the better outside option, it is the low-a workers who pursue the alternative and start firms. (Note that due to the convexity of expected firm value, there will still be some very high level of a above which some high-a workers run firms – but the distribution of a may be very thin there.)

In the contrary case of high ρ, illustrated with $\rho = 0.95$ in the third row of Figure **??**, high-productivity workers have a large edge at running firms. In fact, high ρ not only increases the conditional expectation $\mathsf{E}(\alpha|a) = \mu_\alpha + (\rho\sigma_\alpha/\sigma_a)(a - \mu_a)$ for high-a workers, but also reduces the variance $(1 - \rho^2)\sigma_\alpha^2$ of the distribution of α conditional on a. This implies that the failure probability declines more steeply after its peak and goes to zero more quickly. In the figure, ρ is such that given the distance between $\alpha_R(a)$ and $\mathsf{E}(\alpha|a)$ at high values of a, the success probability for high-a workers is already virtually 1. Low-a workers are not doomed to failure with probability 1 in the case illustrated here, though this can occur for higher ρ that imply steeper $\mathsf{E}(\alpha|a)$.

For intermediate ρ (second row), both high- and low-a workers start firms, whereas people with intermediate a work. In this case, the resulting firm distribution is particularly interesting. In the fourth panel of the second row, the firm size distribution (the kinked line) is plotted. To compare this with the population, the marginal distribution of α in the population is also plotted (top line). To make it comparable with the firm size distribution, the top line does not show $f_\alpha(\alpha)$ as a function of α, but plots it against the labor demand a firm with that α would have. So the x-axis measures firm size. The fraction under the curve that is occupied by the firm size distribution corresponds roughly to the fraction of entrepreneurs in the population. This would be precise if the x-axis was in terms of α.

The firm size distribution consists of two parts, that of firms run by people with high a ($> a_H$), and those with low a ($< a_L$). These are indicated by the thin lines below the firm size distribution. They of course add up to the total distribution. Note that while the bulk of employment is in firms run by high-a people, a significant fraction of the number of

[9] Other parameters as in the top row of Figure 1.

firms is run by people with low a. In this sense, the model can explain why low-productivity firms persist and why productivity heterogeneity is so large: low-productivity firms make low profits, but running them is still optimal for their owners because their outside options is also very low.

With high ρ, these firms are not present, whereas with low ρ, there are only low-a entrepreneurs. Judging from the fact that empirically, there are many small, low-profit firms, but that it has been shown that e.g. successful managers on average also make better entrepreneurs, the intermediate case is the empirically relevant one, putting upper and lower bounds on the correlation between productive and entrepreneurial ability in the population.

With high ρ, another interesting effect becomes visible in the graph: there are more large, high-productivity firms than the marginal distribution of α in the population would suggest. This is because of experimentation and selection: The population of surviving firms is drawn from a distribution of α that is truncated at $\alpha_R(a)$ for each a. The reason is that some entrepreneurs fail because of a bad draw – their first idea was not a good one –, then keep trying, until they draw a sufficiently good idea. Thus, while the population distribution is just the marginal distribution of α, the firm productivity distribution additionally is conditioned on having α above a threshold α_R (see also equation (22)). While this effect occurs for all a, it is most visible here because in the high-ρ case, the entrepreneurs mostly tend to be people with a high outside option, leading to a firm productivity distribution truncated at a high level. For low ρ, the truncation level of the low-a entrepreneurs is much lower, so the experimentation and selection effect is not visible in the figure. Hence, selection is strongest when entrepreneurs have a high outside option, and this occurs when the best entrepreneurs are also very able workers.

7 Unemployment and entrepreneurship

In this setting, labor market frictions do not only cause unemployment, but also influence the occupational choice decision. This is clear from the expression for expected firm value given in equation (16): the possibility of failure, and of being unemployed after it instead of employed as before starting the firm, reduces the incentive to become an entrepreneur.

Consider for instance a reduction in the value of employment because of an additional friction that lengthens unemployment spells, such as firing costs. This is not easy to incorporate into the model with multiworker firms, yet the effect is clear: The reduction in the value of unemployment also reduces the expected value of starting the firm, particularly for people facing a high failure probability. Moreover, it lowers the reservation productivity, worsening the distribution of those firms that are actually around. Lower expected firm value implies that the entry thresholds shift out, and there is less entry for given θ and $\bar{a}$. Whether a_L or a_H reacts more is not clear a priori. Whereas high-a workers have more to lose (the difference $W-U$ increases in a), they also face a lower failure probability at a_H compared to a_L, as shown above.

7.1 An extension: entry by the unemployed

Additional differences arise once unemployed people can also start firms. Clearly, when unemployed and employed people both can start firms, employees will only do so if they are better at it. The reason is that the value of employment exceeds that of unemployment for all a, so at every a where a worker would start a firm, an unemployed person would do so, too. But then there would be no reason to search for employment in the first place. So to obtain situations where both unemployed and employed workers start firms, assume that workers, with their work experience, learn something about how to run a firm, and therefore their productivity draws α are distributed $N(\mu_\alpha + h, \sigma_\alpha^2)$, $h > 0$. This implies a higher expected value of starting a firm for each a, as shown in Figure 3. They face the same productivity threshold α_R for each a, but due to the better distribution of α have a higher probability of success, and in the case of success, better expected outcomes.

As people enter the labor force unemployed, the unemployment-entrepreneurship decision now is the first one to take. As in the case of workers, this defines thresholds a_H^U and possibly a_L^U that partition the population. People with a between these thresholds search for a job, and once they have it, decide whether to keep it or start a firm, with higher expected productivity. If however h is small, the higher outside option of employment implies thresholds $a_L^W < a_L^U$ and $a_H^W > a_H^U$. As all workers have $a_L^U < a < a_H^U$, none of them will start a firm, and only the unemployed start firms. Since this does not fit well with the empirical evidence, h empirically must be high, so that workers face a much higher expected value of firm creation than the unemployed. If it is high enough, workers' thresholds can fulfill $a_L^W > a_L^U$ and $a_H^W < a_H^U$ despite the higher outside option. (This is the case shown in Figure 3.) Then workers with $a_L^U < a < a_L^W$ or $a_H^W < a < a_H^U$ start firms out of employment. As a result, the population is partitioned as follows: very low-ability workers with $a < a_L^U$ start firms directly after entering the labor force, just as workers with very high ability $a > a_H^U$. Workers with slightly low or slightly high ability $a_L^U < a < a_L^W$ or $a_H^W < a < a_H^U$ start firms out of employment, while medium-ability workers with $a_L^W < a < a_H^W$ always prefer to work.[10]

Apart from some more realism, this extension is interesting mainly for considering the

[10]Note that in a full analysis of this case, it should be taken into account that the option of starting a firm enters the value of employment in the problem of unemployed workers deciding between search and entrepreneurship, that is, the simplifying assumption adopted above should be dropped.

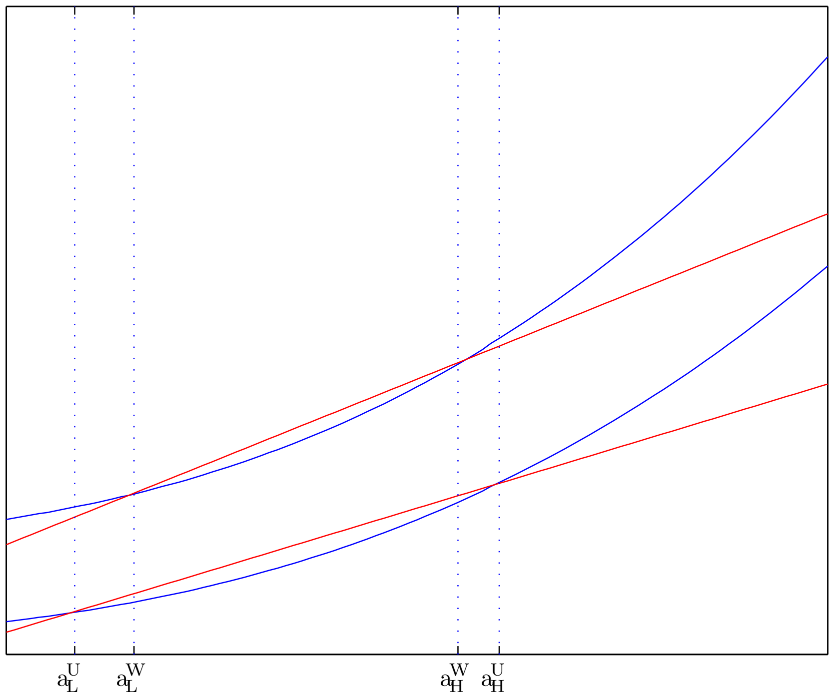

Figure 3: The occupational choice decision for both workers and the unemployed

effect of labor market institutions. The reason is that the job finding probability affects the unemployed much more directly than currently employed workers, and this affects the entry process. Decreases in θq shift the U and W lines in Figure 3 down (they fall for all a), and make them less steep, both effects being stronger for the U line. The expected firm value curves also shift down due to lower value in the case of failure. They shift less than the U line (for $d < 1$), but more than the W line (for $d > \lambda/(1+r)$; both reasonable conditions, remember that λ, the retirement probability, is a very small number). As a result, a_L^U increases and a_H^U falls, while the equivalent thresholds for W move in the opposite direction. So longer expected duration of unemployment, by reducing the value of unemployment, prompts more of the unemployed to try their luck as entrepreneurs. Employees, in contrast, become more loth to risk their jobs in ventures of uncertain success, so less of them start firms.

If the lower job finding probability is due to firing costs, this also reduces the value of success with a firm. As a result, a_L^W falls even more, and a_H^W rises even more. The shifts of the equivalent thresholds for U are now ambiguous, but the result remains that the composition of entrepreneurs worsens as less advantage is taken of workers' experience, and a larger fraction of entrepreneurs is of the least productive type, with $a < a_L^U$. So while the model reproduces the result of a model without occupational choice that firing costs reduce the value of entry and may therefore reduce entry, it generalizes it by recognizing that potential entrepreneurs' outside options are also affected. For workers, the effect is to discourage firm formation, while it may be encouraged for the unemployed. So the main effect of firing costs here is not so much a reduction in entry as a worsening of its composition.

Consequences for the aggregate could be substantial. Remember that the advantage of workers in starting firms needs to be substantial to induce them to do so, as they do in practice. Indeed, this fits with evidence that the unemployed make much less successful entrepreneurs. (There are also previously unemployed entrepreneurs with very high ability that should drive this statistic in the other direction, but maybe a_H^U is very high empirically, so that there are not many people with high enough ability.) Features of the labor market can hence have an impact on the composition of firm founders, and thereby on aggregate productivity.

7.2 A glance at the data

Something akin to the effect of unemployment duration on a_L^U can indeed be observed in the data: Figures 4 and 5 plot country-level indicators of firing costs – known to increase the duration of unemployment both theoretically and empirically – against the fraction of entrepreneurs that state that they have started their venture out of necessity, and not to grasp an opportunity. Firing cost indicators for 2004 are taken from the Fraser Institute Economic Freedom of the World 2006 Annual Report by Gwartney and Lawson (2006). Higher values indicate lower firing costs. Data on entrepreneurship is from the New Entrepreneurship International dataset belonging to the Global Entrepreneurship Monitor (GEM) survey.[11] It covers 46 developed and middle-income countries; I use data for 25 of them for which data was available for at least 4 years in the 2001 to 2005 period. The measure used is the time average of the ratio of two variables, the Necessity Entrepreneurial Activity Index (the number of people involved in total entrepreneurial activity out of necessity relative to the adult population of 18-64 years), divided by the Total Entrepreneurial Activity Index (the number of people currently setting up a business or owning/managing a business existing up to 3.5 years relative to the adult population of 18-64 years). The first group is a subset of the second. The exact question asked is "Are you involved in this start-up to take advantage

[11] Unfortunately, good data on entrepreneurs, particularly for cross-country comparisons, is hard to come by. This survey is one of the most extensive publicly available data sources. Here, the macro overview data has been used, downloaded from EIM's Public Knowledge Web on SMEs and Entrepreneurship at `http://www.entrepreneurship-sme.eu/`. EIM is a Dutch research centre on entrepreneurship, the GEM a global survey run in a collaboration of many research institutes.

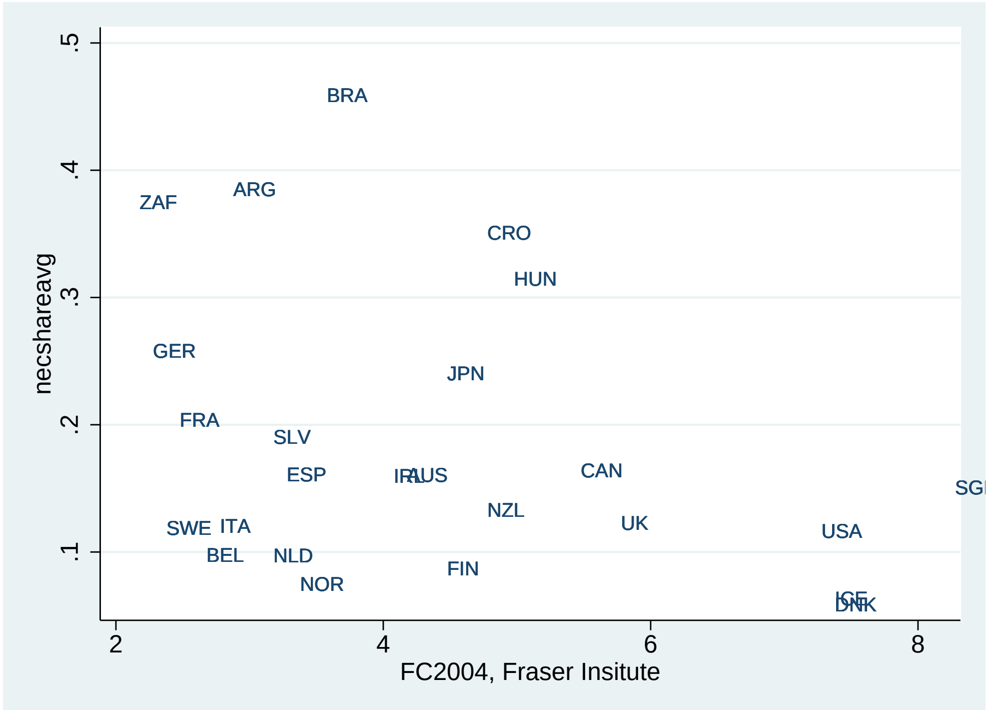

Figure 4: Firing costs versus the share of entrepreneurs who made their choice "out of necessity", 25 countries

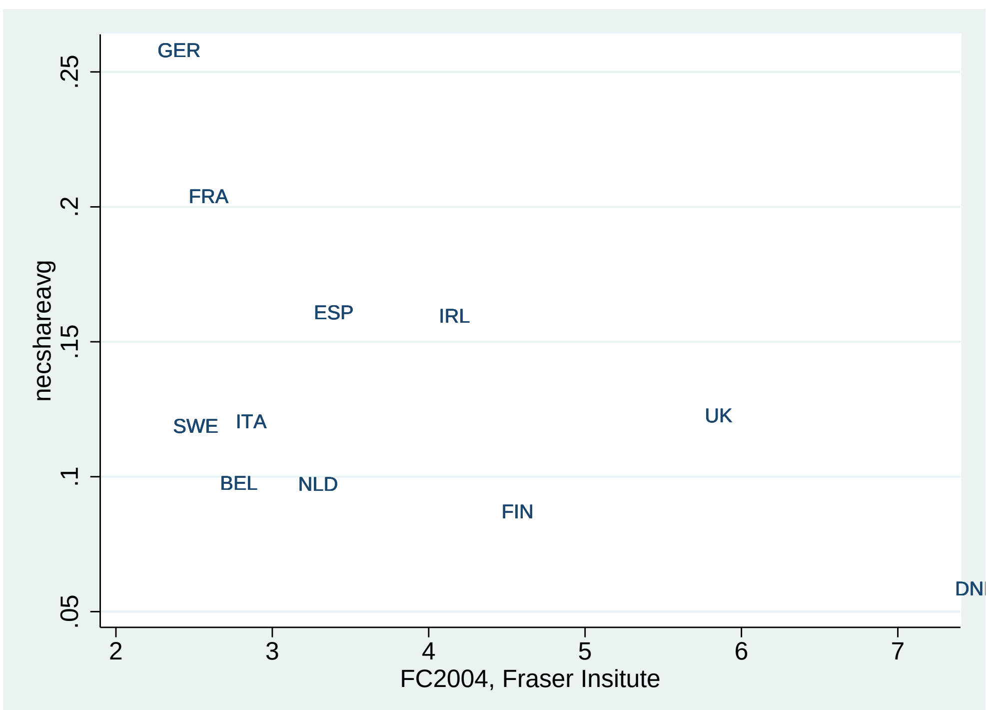

Figure 5: Firing costs versus the share of entrepreneurs who made their choice "out of necessity", 11 EU countries

of a business opportunity or because you have no better choices for work?"

From the figures, a negative correlation between entrepreneurship out of necessity and low firing costs, or a positive correlation with firing costs, is evident both for some EU countries and for a larger and more diverse set of countries. It is striking that a significant fraction of entrepreneurs claim to have made their choice for lack of better alternatives. The cross-country average lies above 10%, with values above 20% in France and Germany, and even higher in some transition economies and Latin American countries. This figure could be argued to correspond to the fraction of entrepreneurs with ability below the a_L thresholds. (Unfortunately, no more detailed breakup of the data is available in the aggregate GEM data; knowing their previous employment status would be useful. However, the associated survey concludes that discontent with the work situation is one of the main drivers for entrepreneurs, and uncertainty avoidance one of the main obstacles.) While a simple correlation of 25 data points does not yet make solid evidence for the model presented here, both the presence and number of entrepreneurs out of necessity and the correlation with firing costs do give it some support.

8 Concluding remarks

This paper has shown that in a reasonably general setting, features of the labor market can affect the decision to become an entrepreneur not only by affecting the firm's value and employment policy in the case of success, but by affecting the prospective entrepreneur's payoff in the case of failure, and by affecting the value of the outside option. This has repercussions for the firm size distribution and for aggregate productivity. The model also illustrated how the correlation of productive ability and entrepreneurial potential in the population shapes the firm size distribution, suggesting that the empirical correlation is positive but not too close to one. It can also explain the existence and persistence of small, low-productivity firms by their owners' low outside option.

Quantifying these results would be interesting. For instance, how large is the effect of labor market regulation on productivity via occupational choice and the composition of entrepreneurs compared to the well-known effect on allocative efficiency of employment? Can the model approximate the empirical firm size distribution? These will be the subject of further work.

Throughout the paper, two related abstractions have been made. Firstly, entry is subject only to the payment of some fixed entry cost which can be financed without frictions. Secondly, if it is optimal for them do so, people can always start a firm if they are employed (in the main setting; in the extension they can always do so). This gives rise to rather extreme life cycles, where some individuals are always workers (or sometimes unemployed), while others start a firm every time they find employment, and never remain employed more than one period. While serial self-employment is a feature of the data, this result seems a bit extreme; the need of time to generate an idea and set up a business plan, or to save to overcome financial frictions, seem plausible remedies, connecting the model to the rich literature on entrepreneurship and liquidity constraints (see e.g. Evans and Jovanovic 1989, Cagetti and De Nardi 2006).

The contribution of the paper, however, is independent of these abstractions; it is to go beyond the impact of labor market circumstances on firm employment or individual search and savings decisions, but to link the two sides through the occupational choice decision. This has broad implications, for instance, employment protection legislation could have much richer effects on productivity and the firm size distribution than when it only affects labor demand. Policy implications also follow; entry subsidies for instance may end up financing mainly low-productivity firms, particularly when labor markets are rigid. This is particularly true for start-up subsidies to the unemployed. These can hence not be evaluated independently from the situation on the labor market.

References

Acemoglu, D. and Hawkins, W. (2006), 'Equilibrium Unemployment in a Generalized Search Model', *mimeo* .

Bartelsman, E. J., Haltiwanger, J. and Scarpetta, S. (2009), 'BHS05: Measuring and analyzing cross-country differences in firm dynamics', *Producer Dynamics: New Evidence from Micro Data* p. 15.

Bertola, G. (2004), 'A Pure Theory of Job Security and Labour Income Risk', *Review of Economic Studies* **71**(1), 43–61.

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

Cahuc, P., Marque, F. and Wasmer, E. (2004), 'A Theory of Wages and Labor Demand with Intrafirm Bargaining and Matching Frictions', *CEPR Discussion Paper* **4605**.

Caves, R. E. (1998), 'Industrial Organization and New Findings on the Turnover and Mobility of Firms', *Journal of Economic Literature* **36**(4), 1947–1982.

Evans, D. S. and Jovanovic, B. (1989), 'An Estimated Model of Entrepreneurial Choice under Liquidity Constraints', *Journal of Political Economy* **97**(4), 808–827.

Felbermayr, G. and Prat, J. (forthcoming), 'Product Market Regulation, Firm Selection and Unemployment', *Journal of the European Economic Association* .

Fonseca, R., Lopez-Garcia, P. and Pissarides, C. A. (2001), 'Entrepreneurship, start-up costs and employment', *European Economic Review* **45**(4-6), 692–705.

Foster, L., Haltiwanger, J. and Krizan, C. J. (2001), Aggregate Productivity Growth: Lessons from Microeconomic Evidence, *in* C. R. Hulten, E. R. Dean and M. J. Harper, eds, 'New Developments in Productivity Analysis', NBER Studies in Income and Wealth, University of Chicago Press, Chicago.

Gwartney, J. D. and Lawson, R. A. (2006), *Economic Freedom of the World: 2006 Annual Report*, Fraser Institute, Vancouver.

Hintermaier, T. and Steinberger, T. (2005), 'Occupational choice and the private equity premium puzzle', *Journal of Economic Dynamics and Control* **29**, 1765–1783.

Hopenhayn, H. (1992), 'Entry, Exit, and Firm Dynamics in Long Run Equilibrium', *Econometrica* **60**(5), 1127–1150.

Hopenhayn, H. and Rogerson, R. (1993), 'Job Turnover and Policy Evaluation: A General Equilibrium Analysis', *Journal of Political Economy* **101**(5), 915–938.

Lazear, E. P. (2004), 'Balanced Skills and Entrepreneurship', *American Economic Review Papers and Proceedings* **94**(2), 208–211.

Ljungqvist, L. (2002), 'How Do Lay-Off Costs Affect Employment?', *The Economic Journal* **112**(482), 829–853.

Luttmer, E. (2007), 'Selection, Growth and the Size Distribution of Firms', *Quarterly Journal of Economics* **122**(3), 1103–1144.

Miller, R. A. (1984), 'Job Matching and Occupational Choice', *Journal of Political Economy* **92**(6), 1086–1120.

Moskowitz, T. J. and Vissing-Jørgensen, A. (2002), 'The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?', *American Economic Review* **92**(4), 745–778.

Pissarides, C. A. (2000), *Equilibrium Unemployment Theory*, 2nd edn, MIT Press, Cambridge, MA.

Poschke, M. (2008), 'Who Becomes an Entrepreneur? Labor Market Prospects and Occupational Choice', *IZA Discussion Paper* **3816**.

Santarelli, E. and Vivarelli, M. (2006), 'Entrepreneurship and the Process of Firms' Entry, Survival and Growth', *IZA Discussion Paper* **2475**.

Stole, L. A. and Zwiebel, J. (1996), 'Organizational Design and Technology Choice under Intrafirm Bargaining', *American Economic Review* **86**(1), 195–222.

Tybout, J. R. (2000), 'Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?', *Journal of Economic Literature* **38**(1), 11–44.

A Derivation of the wage equation

The wage is given by the linear ODE

$$\bar{w} = (1-\eta)\frac{r}{r+\eta\lambda}rU + \eta\frac{r+\lambda}{r+\eta\lambda}(\bar{a}y'(n) - \bar{n}\bar{w}'(\bar{n})). \tag{24}$$

Without the constant, this is

$$\bar{w}'(\bar{n})) + \frac{\bar{w}}{c_1 \bar{n}} - \frac{\bar{a} y'(n)}{\bar{n}} = 0, \tag{25}$$

where $c_1 = \eta(r + \lambda)/(r + \eta\lambda)$. The solution of the homogeneous equation

$$\bar{w}'(\bar{n}) + \frac{\bar{w}}{c_1 \bar{n}} = 0 \tag{26}$$

then is

$$\bar{w}(\bar{n}) = C\bar{n}^{-1/c3}. \tag{27}$$

C is a constant of integration that can be a function of $\bar{n}$. So take the derivative of equation (27) with respect to $\bar{n}$:

$$\frac{\partial \bar{w}}{\partial \bar{n}} = C'(\bar{n})\bar{n}^{-1/c_1} - \frac{C}{c_1}\bar{n}^{-1/c_1 - 1} \tag{28}$$

Substituting this into (25) yields

$$C'(\bar{n}) = \bar{a} y'(n) \bar{n}^{1/c_1 - 1}. \tag{29}$$

Integrating this gives $C(\bar{n})$ as

$$C(\bar{n}) = \int_0^{\bar{n}} \bar{a} y'(\bar{a}z) z^{1/c_1 - 1} \mathrm{d}z + D, \tag{30}$$

so the wage $\bar{w}$ is

$$\bar{w}(\bar{n}) = \bar{n}^{-1/c_1} \int_0^{\bar{n}} \bar{a} y'(\bar{a}z) z^{1/c_1 - 1} \mathrm{d}z + D\bar{n}^{-1/c3}. \tag{31}$$

The constant D can be dealt with by assuming that the wage bill goes to zero as employment goes to zero; this implies $D = 0$. The solution to equation (24) then is

$$\bar{w}(\bar{n}) = \bar{n}^{-1/c_1} \int_0^{\bar{n}} \bar{a} y'(\bar{a}z) z^{1/c_1 - 1} \mathrm{d}z + \frac{r(1-\eta)}{r + \eta\lambda} rU.^{12} \tag{32}$$

Integrating the integral by parts (integrate the term in $\bar{n}$ and differentiate the marginal product), substituting in the production function (1) and using $n = \bar{n}\bar{a}$ then yields

$$\bar{w}(\bar{n}) = \frac{r(1-\eta)}{r + \eta\lambda} rU + c_1 \bar{a} y'(n) - \frac{c_1 \gamma(\gamma - 1)\bar{a}^{\gamma}\bar{n}^{\gamma - 1}}{\gamma - 1 + 1/c_1}. \tag{33}$$

[12] Up to here, the argument follows Cahuc et al. (2004) quite closely.

The last term here comes from the overhiring effect. To obtain the wage at the firm's desired employment as given by the labor demand condition (8), take the derivative with respect to $\bar{n}$ and multiply by $\bar{n}$ to replace the $\bar{n}\bar{w}(\bar{n})$ term in (8). This gives the labor demand condition

$$\bar{a}y'(n) = \frac{k_v}{q} + \bar{w} + \frac{\gamma(\gamma - 1)\bar{a}^{\gamma}\bar{n}^{\gamma-1}}{\gamma - 1 + 1/c_1}. \tag{34}$$

Solving for the last term and substituting into (32) then yields the wage for an effective worker at the optimal employment level as

$$\bar{w} = \frac{\eta(r + \lambda)}{r(1 - \eta)} \frac{k_v}{q} + rU. \tag{35}$$