February 1999

THE INFORMATIONAL CONTENT OF THE BUSINESS CYCLE*

Harris Dellas
Department of Economics
University of Bern

Abstract

Economic decisions -occupational and entrepreneurial choices- may violate true comparative advantage when economic agents are uncertain about which activity best matches their talents. If relative performance varies over the business cycle (for instance, if downturns affect disproportionately those who are pursuing the wrong activity), then economic fluctuations may affect the probability and persistence of resource mismatches. The present work offers a novel, informational perspective to the business cycle and provides a link between aggregate fluctuations and the efficient allocation of resources.

JEL classification: D83, E32, E60
Key words: Informational content of the business cycle, matching

VWI, University of Bern, Gesellschaftsstrasse 49, CH-3012 Bern, Switzerland, email: harris.dellas@vwi.unibe.ch

*An earlier version of this paper -circulated under the title "Reccessions and ability discrimination"- benefited from conversations with P. Aghion, R. Flood, K. Rogoff, J. Rotemberg, M. Woodford and T. Zha.

Economic decisions are trivial when individual talents and relative payoffs to alternative activities are fully known. In the absence of increasing returns each resource exploits its comparative advantage. In reality, however, both talents and payoffs are subject to considerable uncertainty. Individuals seldom know in advance which opportunities (activities) best match their abilities. Training programs (for example, the formal educational system) and other experiences can be useful for testing one's relative and absolute abilities but they are far from fully revealing and can at best provide only an indication of future performance. It usually takes a prolonged period of practice in the selected line of business to form a knowledgeable assessment of one's qualities. Moreover, in a rapidly changing economic environment, the process of "learning" about the most profitable matches may have to be repeated. Consequently the lack of complete information about the best association between personal talents and existing opportunities can produce mistakes in the selection of economic activities and lead to allocations that violate true comparative advantage. An extremely important question then regards the existence of mechanisms that improve the effectiveness of the selection process. This paper identifies the business cycle as a sorting mechanism that can limit the possibility and duration of resource misallocation. Recessions seem a priori relevant because they represent economic adversity; and it is well known from the natural world that adversity plays an important role in revealing one's level of fitness.

The—- thesis that economic fluctuations may be intrinsically related to economic efficiency and long term prospects was popular among the pre Great Depression era economists. Schumpeter (1991), in particular, viewed business cycles as a manifestation of the evolutionary process of innovation, as a reflection of the replacement of the old (and inefficient) by the new (and efficient). Such views, however, were pushed to the side by the experience of the Great Depression and the rooting of the Keynesian revolution which preached the benefits of macroeconomic stability (viewing business cycles as the outcome of easily correctable market failure). Only recently have there been some attempts to re-evaluate the role of the business cycle.

In this paper we argue that the business cycle can influence efficiency in the allocation of resources by affecting the speed and accuracy of learning about one's best match. A simple example can illustrate how this works. Consider a primitive society which engages in two types of activities, hunting and farming. Let the members of this community come -genetically- in two types; those who have an absolute advantage in hunting; and those who have an advantage in farming. Furthermore, assume that the individuals do not know their true type but know the probability distribution of performance in hunting and farming for each type as a function of the state of nature. In each period, nature randomly dictates the availability of prey in hunting and of farm products in farming. Consider now those starting out their careers as hunters. After repeated sampling, each new hunter will be able to -perhaps imperfectly in a finite sample- infer his type and make an occupational choice[1]. If the two types are not perfectly distinguishable then some people may misclassify themselves. The misclassification error will depend on the sample size and on the difference in the parameters of the probability distributions of performance across the two types; the less the two types are alike in terms of ability, the faster the revelation of type and the smaller the likelihood of an occupational mistake. The error may also depend on the favorableness of the state of nature during sampling. It is conceivable that when prey is plentiful, the performance of both types will be comparable (shooting accuracy is not a big factor when one runs into a large herd of animals). On the other hand, when prey is scarce, the "natural" hunters may significantly outperform the rest. Adverse conditions can thus prove more informative, and may have the potential to improve the selection process and the allocation of resources. A special case of interest is when severe adversity can immediately induce a separating equilibrium[2].

Similar considerations are at work in many economic decisions, and in particular, in those pertaining to occupational and entrepreneurial choices. It seems plausible then that recessions may contribute to greater economic efficiency by providing people with the information they need in order to make more knowledgeable economic decisions. This does *not*, however, mean that similar informational advantages will be enjoyed by an economy which experiences economic fluctuations compared to one that successfully stabilizes both sides of the business cycle. This is due to the fact that expansions may carry too little information for discrimination. Whether the net sum over the business cycle is greater than that achieved under some or perfect stabilization depends on several things: The form of skewness in comparative -across types- cyclical performance; the symmetry of the effects of business cycle across activities; and the type of stabilization. We show that the curvature of the function of relative cyclical performance plays a critical role. In general, the exercise of stabilization policy creates informational inefficiencies when the difference in the performance of different types is a decreasing and convex function of the favorableness of the state of the business cycle[3].

The model has implications for, among other things, the pattern of firm dissolutions and the properties of the empirical distribution of performance in any activity. These predictions can be used to infer the shape of the function of relative performance and hence to assess the informational role of recessions and/or the business cycle. In the empirical section we examine the empirical distribution of growth rates of sales of the firms contained in the Compustat data base. The -implied- shape of the function of relative performance indicates that recessions can enhance the ability to discriminate among types. The analysis, however, is less successful in establishing the net informational contribution of the business cycle.

Section 1 describes the optimal sequential procedure used in occupational (entrepreneurial) decisions. Section 2 derives the empirical implications and Section 3 presents the empirical findings.

1.The Theory

Consider an economy that consists of two types of individuals: h-type (high ability) and l-type (low ability); and permits two possible activities (careers): entrepreneurship (denoted by e) and workmanship (denoted by w). Let each type be "best match" to one of the two activities (the concept of "best match" will be made more precise shortly). Assume also that one's true type is not known with certainty but may be revealed - partially or fully- from observations of performance. Let x_{jt} be the performance of an individual -who does not know his true type i, $i = h, l$- in activity j, $j = e, w$, in period t. Let x_{jt} depend, in addition to one's type on two exogenous, random variables, s_{jt} and

u_{ijt}; s_{jt} is a contemporaneously observable variable representing the favorableness of the aggregate state of nature in activity j, which, in our model is captured by the state of the business cycle; u_{ijt} is unobservable and represents the effects of luck. We will assume that s and u are uncorrelated and that the *variance* of u is independent of s (the latter assumption is critical for the main results of this paper). The agents' objective is to device an optimal lifetime activity -career- plan.

The occupational choice faced during any period can be most fruitfully analyzed as a sequential statistical game with the economic agent playing a game against nature. In each period, the agent must select an activity to participate in. Nature then moves and generates exogenously an observation of performance, x_j. Such a decision process falls in the category of two-armed bandit games. In the present paper we will focus on a special case of this class of statistical games by imposing some restrictions on the menu of actions that will be available at any point in time. In particular, we will assume that after the initial choice of activity -which occurs before any observations have been taken- each individual is allowed only one career switch in his lifetime. This does not represent any loss of generality and can be justified by appealing to appropriately large fixed costs that may result from multiple career changes[4].

Assume that the random variables $x_{j1}, x_{j2}, \ldots,\ j = e, w$ representing performance in activity j are available for observation and let $f(x_{jn}|i), n = 1, 2, ...$ be the corresponding density functions. Given the subject of this work, it is natural to permit the x's to be taken sequentially. Let Z_t be the sample space of x_{jt}, $t = 1, 2, \ldots$ and assume that $X_t = (x_{j1}, x_{j2}, .., x_{jt}\}$, has density $\phi(X_t|i), i = h, l$, on $Z_1 \times .. \times Z_t$. Attention will be restricted to situations in which the x's are independent observations so that $\phi(X_t|i) = \Pi_{n=1}^{t} f(x_n|i)$.

Our objective is to derive and characterize the optimal sequential decision procedure. We will assume that the choice of activity to be pursued in period t, $t = 1, 2, \ldots$, is made before the current observation becomes available. That is, the decision is conditional on a sample consisting of observations extending up to the previous period, $t - 1$. The optimal procedure essentially involves two decisions: the initial choice of activity made in the beginning of period 1; and the subsequent decision to switch or not made in period

t, $t > 1$.

At each point in time, an individual[5] can be thought as testing two hypotheses: H_{et} : better off being in entrepreneurship in period t, against H_{wt}: better off being in workmanship in period t. Let us consider the problem faced by someone who finds himself in entrepreneurship in the beginning of period t, $t = 2, 3, ...$ An activity is selected in order to maximize expected lifetime income,

$$\begin{aligned} E_t V_e(p_t, s_t, u_{et}) = \max\{ & E_t V_w(p_t, s_t, u_{wt}), \\ & E_t[Y_e(p_t, s_t, u_{et}) + \beta V_e(p_{t+1}, s_{t+1}, u_{et+1})]\} \end{aligned} \tag{1}$$

subject to the law of motion of p_{t+1}

$$p_{t+1}(h) = \frac{p(h)P(X_t|h)}{p(h)P(X_t|h) + p(l)P(X_t|l)} \tag{2}$$

where $p_{t+1}(h)$ is the posterior probability of being an h-type after having observed X_t, $p(i)$ is the prior probability[6] of being an i-type and $P(X_t|i)$ is the likelihood function of X for type i. V_j is the value function of an individual operating in activity j, $\beta \epsilon [0, 1]$ is the discount factor, E_t is expectation conditional on p_t and

$$u_e = \{u_{he}, u_{le}\}, u_w = \{u_{hw}, u_{lw}\}$$

$$E_t V_e(p_t, s_t, u_{et}) = p_t V_{he}(s_t, u_{het}) + (1 - p_t)V_{le}(s_t, u_{let}) \tag{1a}$$

$$E_t Y_e(p_t, s_t, u_{et}) = p_t E_t x_{he}(s_t, u_{het}) + (1 - p_t)E_t x_{le}(s_t, u_{let}) \tag{1b}$$

$$E_t V_w(p_t, s_t, u_{wt}) = \max\{E_t[Y_w(p_t, s_t, u_{wt}) + \beta V_w(p_{t+1}, s_{t+1}, u_{wt+1})]\} \tag{1c}$$

$$E_t V_w(p_t, s_t, u_{wt}) = p_t V_{hw}(s_t, u_{hwt}) + (1 - p_t)V_{lw}(s_t, u_{lwt}) \tag{1d}$$

$$E_t Y_w(p_t, s_t, u_{wt}) = p_t E_t x_{hw}(s_t, u_{het}) + (1 - p_t)E_t x_{lw}(s_t, u_{let}) \tag{1e}$$

where V_{ij} is the value function of type i in activity j, and x_{ij} is the performance of type i in j. Equations (1a)-(1e) define the expected payoffs associated with each type in each activity.

The RHS of (1) consists of two terms. The first one is the expected lifetime earnings of an entrepreneur who switches to workmanship. The second term is the expected payoff enjoyed by someone who decides to remain for at least one more period in entrepreneurship. Note that the decision problem of an agent whose initial choice was workmanship is completely analogous to that described in (1)). For the occupational decision to be nontrivial each type must be better off in a different activity. We will assume -Assumption 1- that the high type's best match is in entrepreneurship and the low type's in workmanship (this assignment is also assumed to be independent of the degree of stabilization).

Assumption 1 (Best match): $E_t x_{he} > E_t x_{hw}, E_t x_{lw} > E_t x_{le}$

We first prove the existence of an interior solution for the choice of activity (Proposition 1). We then show that such an equilibrium can be reached from any set of prior beliefs (Proposition 2). And, finally, we characterize the properties of the equilibrium as a function of the degree of stabilization policy (Propositions $4, 5, 6$).

Proposition 1 *There exists a* $\tilde{p}(X_t)$:

$$E_t[Y_e(p_t, s_t, u_{et}) + \beta V_e(p_{t+1}, s_{t+1}, u_{et+1})] \geq E_t[V_w(p_t, s_t, u_{wt}) \ for \ p_t \geq \tilde{p}$$
$$E_t[Y_e(p_t, s_t, u_{et}) + \beta V_e(p_{t+1}, s_{t+1}, u_{et+1})] < E_t[V_w(p_t, s_t, u_{wt}) \ for \ p_t < \tilde{p}$$

The optimal rule is to remain in entrepreneurship as long as $p_t \geq \tilde{p}$ and to switch to workmanship otherwise.

We now derive the optimal decision using the Sequential Probability Ratio Test, $SPRT$ (Wald, 1947). According to Proposition 1, *workmanship* is the activity of choice in period $t, \ t > 1$ if

$$\begin{gathered} p_t(h|X_{t-1})E_t[Y_{\text{het}} + \beta p_{t+1}(h|X_t)V_{\text{het}}] + \\ p_t(l|X_{t-1})E_t[Y_{\text{let}} + \beta p_{t+1}(l|X_t)V_{\text{let}}] \\ < \\ p_t(h|X_{t-1})E_t[Y_{\text{hwt}} + \beta p_{t+1}(h|X_t)V_{\text{hwt}}] + \\ p_t(l|X_{t-1})E_t[Y_{\text{lwt}} + \beta p_{t+1}(l|X_t)V_{\text{lwt}}] \end{gathered} \tag{3}$$

which can be rewritten as

$$\frac{p_t(h|X_{t-1})}{p_t(l|X_{t-1})} < \frac{(W_{lwt} - W_{let})}{(W_{het} - W_{hwt})} \equiv A_t \tag{4}$$

where $W_{het} = E_t[Y_{\text{het}} + \beta p_{t+1}(h|X_t)V_{\text{het}}$ and so on. Then W_{ijt} is the expected lifetime earnings of type i who stayed in activity j during period t. Hence, $W_{lwt} - W_{let}$ is the loss suffered by a low type who decides to spend an additional period in entrepreneurship; and $W_{het} - W_{hwt}$ is the loss suffered by a high type who permanently switches to workmanship. Using Bayes' rule in (4) we have

$$\frac{p_t(X_{t-1}|h)}{p_t(X_{t-1}|l)} < A_t \frac{p(l)}{p(h)} \equiv R_t \tag{5}$$

Let $z_n = \log[f(x_n|h)/f(x_n|l)]$ and $S_t = \Sigma_{n=1}^t z_n$. Taking logarithms in equation (5) produces Proposition 2.

Proposition 2 *The optimal decision of an entrepreneur*[7] *in period* t, $t > 1$, *is: stay in entrepreneurship if* $S_{t-1} \geq \ln R_t$;*otherwise, switch to workmanship.*

Using the definition of the posterior probability distribution, the criterion can be restated in terms of the posterior probability,

$$If\ p_t \geq A_t/(1 + A_t) \text{ then stay in entrepreneurship;}$$
$$\text{otherwise switch to workmanship.}$$

Hence, the critical value of p, $\tilde{p}$, is given by

$$\tilde{p} = A_t/(1 + A_t) \tag{6}$$

The decision problem faced in the beginning of period 1 (the initial selection) is somewhat different from that faced during subsequent periods. This is due to the fact that the initial decision is always reversible. Consequently, the expected earnings of a high type who starts out in workmanship exceed those of a high type who -mistakenly- switches from entrepreneurship to workmanship ($V_{hwt} < V_{hw1}, t > 1$) due to the value of the option to switch. Nevertheless, the criterion for the initial selection takes the same form as before, namely,

$$\begin{gathered}\text{If } p_t \geq A_{1t}/(1+A_{1t}) \text{ select entrepreneurship;} \\ \text{otherwise select workmanship}\end{gathered} \tag{7}$$

where A_{1t} is defined by equations analogous to those defining A_t. It can be seen that $A_1 > A$, which implies that the critical value $\tilde{p}(A_1)$ exceeds $\tilde{p}(A)$. The possibility of an irreversible occupational error dictates greater caution.

It remains to show that an interior solution for the allocation of resources exists for all sets of prior beliefs. We need to prove that the sequential procedure described above is proper -that is, it terminates with probability one- for a low type sampling in entrepreneurship (and a high type sampling in workmanship). Let N_{le} be the stopping (switch) time of the $SPRT$ of the low type whose initial choice was to sample in entrepreneurship; and P_{le} be the *pdf* of N_{le}. If $f(x_t|h)$ and $f(x_t|l)$ differ with positive probability -which is required if the definition of type is to be meaningful- then it is well known that (see, for instance, Berger, 1985)

Proposition 3 $P_{le}(N_{le} < \infty) = 1, P_{hw}(N_{hw} < \infty) = 1$

Proposition 3 implies that mismatches arising from the initial selection of activity are eventually corrected. Mismatches due to mistaken switching, however, are permanent (because of the single switch assumption).

To facilitate the exposition we will now make the assumption that the observations of performance come from normal distributions. Namely,

$$x_{jt} \sim N(\mu_{ij}(s_{jt}), \sigma^2) \qquad i = h, l \quad j = e, w$$

We will set $s_{et} = s_{wt}$ that is, we will focus on aggregate shocks exclusively.

Normality is helpful for calculating the estimator of type. In general, the form of the optimal estimator depends on the form of the loss function used. However, under normality the posterior distribution is symmetric and thus the posterior mean is the optimal estimator of type -in our case the mean- independent of the form of the loss function[8]. Let

$X^*_{ie} = n^{-1}\Sigma^n_{t=1} x_{iet}$

Consequently,

$X^*_{he} \sim N(\mu^*_{he}, \sigma^2/n)$ if sampled from the h-population

$X^*_{le} \sim N(\mu^*_{le}, \sigma^2/n)$ if sampled from the l-population

where $\mu^*_{he} = n^{-1}\ \Sigma^n_{\tau=1}\mu_{he}(s_\tau)$ and $\mu^*_{le} = n^{-1}\ \Sigma^n_{\tau=1}\mu_{le}(s_\tau)$ and $\mu^*_{he} > \mu^*_{he}$

The $SPRT$ criterion for a high type in period $t > 1$ takes the form

$$\text{Select entrepreneurship if } X^* \geq \xi; \quad \text{otherwise select workmanship} \tag{8}$$

where $\xi = (\mu^*_h + \mu^*_l)/2 - [(\sigma^2/t)/(\mu^*_h - \mu^*_l)]$

lnR

In the special case where both the prior probabilities and the loss functions associated with the wrong decision are equal, the criterion becomes very simple as $\xi = (\mu^*_h + \mu^*_l)/2$. In general, the larger the prior probability of being the high type and/or the larger the cost of mistakenly giving up entrepreneurship the more conclusive the sample evidence in favor of being the low type has to be (a low X^* or equivalently, a low $\tilde{p}$) in order for an individual to be induced to switch to workmanship. A large prior of being a low type together with high payoffs in workmanship relative to entrepreneurship support a high critical value for X^* and $\tilde{p}$.

In the next section we turn to the examination of how policy can affect the speed of learning (type revelation), the average length of sampling (persistence of transitory mismatches) and the probability of a permanent type misclassification. All of these considerations are of obvious importance for evaluating the implications of macroeconomic stabilization policy for short and long term efficiency in the allocation of resources. We study the effects of policies that stabilize the cycle by comparing two regimes. Under the first one, business cycles are allowed to take place (s varies exogenously). Under the other regime, the government stabilizes s at its average value, s^* (the business cycle is completely eliminated[9]). Intermediate cases can be easily accommodated. As a result of the distributional assumption made above, the distance between the population means will play the critical role in the evaluation of the allocation effects of policy. Under the former regime, the distance is given by

$$\mu_h^* - \mu_l^* = n^{-1}\Sigma_{\tau=1}[\mu_h(s_\tau) - \mu_l(s_\tau)] = n^{-1}\Sigma g(s_\tau) \tag{9}$$

and under the latter

$$\mu_h^* - \mu_l^* = \mu_h(s*) - \mu_l(s*) \tag{10}$$

If g is convex (concave), then the RHS of (9) is greater (smaller) than that of (10) -from Jensen's inequality- and the distribution functions for the two types are further apart (closer) in the absence of stabilization policy. We now study several important dimensions of allocation efficiency under alternative policy regimes maintaining the assumption of normality.

1.I The speed of type revelation (learning)

Equation (2) can be written as

$$p_{t+1} = \frac{1}{1 + LF_t[p(l)/p(h)]} \tag{11}$$

where LF_t is the ratio of the likelihood functions for the low and high type respectively. Note that

$$\lim LF_{t(t\to\infty)} = \{^0_\infty \Leftrightarrow \lim p_{t(t\to\infty)} = \{^{1 \text{ if true type is h}}_{0 \text{ if true type is l}}$$

so that true type is revealed with probability one. The value of LF determines the slope of the time path of p_t and hence, it can be loosely thought as measuring the speed of type revelation (learning). The question of interest concerns how stabilization policy affects the slope of the time profile of p.

Proposition 4 *For any sample size t, the average value of LF_t is decreasing in $\mid \mu_h^* - \mu_l^* \mid$ when the true type is h (increasing when true type is l).*

If g in equation (9) is a decreasing function of s then recessions offer the best opportunities for type discrimination. If g is also convex in s then the distance between the means of the probability distributions of the high and low types is greater in the absence of stabilization. Under these conditions stabilization leads to a slower average "speed" of learning about one's type by making the time profile of p flatter.

The function $g(s)$ is decreasing if the fortune of the low ability entrepreneur is more sensitive to the state of the business cycle than that of the high ability one; $g(s)$ is convex if this sensitivity becomes more pronounced as recessions become more severe. In general, convexity means that the difference in relative performance in entrepreneurship between the two types (as measured by variables such as profits, sales, employment losses, etc.) is greater when times for entrepreneurs are bad. If this is the case, the elimination of recessions shuts down a mechanism that carries the most valuable piece of information needed by aspiring businessmen to infer their true type. Learning takes place even when the business cycle is completely stabilized but at a slower rate (the relative contribution of noise tends to increase disproportionately with the favorableness of the business if g is convex and the variance of u is independent of s).

If, on the other hand, relative performance is a concave function of the business cycle then stabilization policy contributes to faster revelation. While one can speculate about the plausibility of alternative shapes this is an issue that can only be settled by the empirical evidence (see section 3).

It must be noted that the assumption made earlier concerning the relationship between the distributions of u and s is absolutely critical for the results obtained. Recessions are more informative when they are associated with a smaller "overlap" of the distributions of the two types. This is true when the average difference in performance between the two types is decreasing in the favorableness of the business cycle *and,* at the same time, the variance of the unobserved idiosyncratic shock is not decreasing in the favorableness of the business cycle. We have assumed that the variance of u is independent of s. Recessions may well be periods of slower learning if they are associated with a sufficiently larger variance for u. While it is hard to think of reasons that would make personal luck more volatile during bad times one cannot rule this possibility out.

Nevertheless, it is obvious how changing this assumption would affect the results.

1.II. The expected stopping time

An important dimension of the allocation of resources concerns the duration of transitory mismatches. We have shown that the process of type revelation is slower under stabilization if g is convex. Slower, however, does not necessarily translate into longer because the position of the finish line, the value of $\tilde{p}$, matters too. We now turn to the calculation of the average stopping (switch) time for a low type sampling in entrepreneurship.

The next theorem, due to Wald (1947), can be used to approximate the expected number of observations, $E(N)$, required by a $SPRT$.

Theorem: Suppose that $z_1, z_2, \ldots$ is a sequence of independently and identically distributed random variables such that $E(z_t) = m, t = 1, 2, \ldots$. For any sequential procedure for which $E(N) < \infty$ the following relation must be satisfied:

$$E(z_1 + z_2 + \ldots + z_N) = mE(N)$$

Proposition 5 *The approximate expected stopping time is given by*

$$E(N) = \frac{E \ln R}{E(z)} \tag{12}$$

The relationship between policy and the persistence of occupational mismatches in entrepreneurship depends on the curvature of g and also on how policy affects relative incomes in the two activities. R is increasing or decreasing in the degree of stabilization depending on whether stabilization benefits the workers more relative to the entrepreneurs. If the effects are symmetric across activities for both types or if policy favors the entrepreneurs, then the stopping time and job mismatches in entrepreneurship are unambiguously prolonged under stabilization. On the other hand, if the workers are the main beneficiaries of stabilization and if this effect is strong, then the stopping time may be shortened as the incentive to become a worker increases. Note, however, that in the latter case sampling in workmanship becomes more prolonged, which means that asymmetry in the effects of the business cycle leads to a trade off in persistence in sampling across activities.

An alternative but equivalent way of studying the persistence of transitory mismatches is through the analysis of the effects of stabilization on the value of the switch probability $\tilde{p}$. Equation (6) implies that the critical value is increasing in A. If policy does not favor the workers then $\tilde{p}(stab) \geq \tilde{p}(no-stab)$; otherwise the inequality is reversed. In the latter case, stabilization implies longer sampling because of its lower "speed" of learning. In the former case, its effect on $E(N)$ is ambiguous: p_t travels a shorter distance at a slower speed.

1.III The probability of misclassification.

We now turn to a related issue, namely, how the exercise of stabilization policy affects the probability of type misclassification. There are two types of misclassification that may be committed by individuals who are sampling in entrepreneurship. One is transitory and the other permanent. The transitory error is associated with the low types who continue practicing entrepreneurship. It was described in the previous sub-section. The permanent one -which arises from the assumption of an irreversible career switch- is associated with those high type agents who mistakenly conclude that their type is low and exit.

Recall that the entrepreneur's criterion for a classification-occupation decision in period t is: exit if $X^* < \xi$; stay if $z^* \geq \xi$. Consequently, the probability of a permanent misclassification is (see Anderson, 1958)

$$\text{Prob}(X^* < \xi|\text{true type} = h) = \text{Prob}[\sqrt{t}(X^* - \mu_h^*)/\sigma < \sqrt{t}(\xi - \mu_h^*)/\sigma|\text{type} = h] =$$

$$= F(\sqrt{t}(\xi - \mu_h^*)/\sigma)$$

and of a transitory misclassification

$$\text{Prob}(X^* > \xi|\text{true type} = l) = \text{Prob}[\sqrt{t}(x^* - \mu_l^*)/\sigma > \sqrt{t}(\xi - \mu_l^*)/\sigma|\text{type} = l] =$$

$$= 1 - F(\sqrt{t}(\xi - \mu_l^*)/\sigma)$$

where F is $N(0, 1)$.

If the loss functions are symmetric and the priors equiprobable $R = Ap(l)/p(h) = 1$ and stabilization policy leads to higher short and long term misclassification errors when g is convex. Its effects, however, are ambiguous if $R \neq 0$, for exactly the same reason that its effects on $\tilde{p}$ were found to be ambiguous. We have the following result.

Proposition 6 *Let g be a convex function of the business cycle. The probability of permanent misclassification is increasing in the degree of stabilization if the exercise of stabilization policy does not favor the entrepreneurs.*

Proof: $(\xi_t - \mu_h^*) = -(\mu_h^* - \mu_l^*)/2 - [(\sigma^2/t)/(\mu_h^* - \mu_l^*)]lnR$. F is increasing. We have shown that $lnR > 0$ (otherwise the individual would not be in entrepreneurship). Whether $\xi_t - \mu_h^*$ is decreasing in $\mu_h^* - \mu_l^*$ or not depends on $B = sign\ \{dlnR/d[\mu_h^* - \mu_l^*]\}$. If stabilization benefits workers at least as much as it benefits entrepreneurs then $B > 0$. Consequently, $\xi_t - \mu_h^*$ is decreasing in $\mu_h^* - \mu_l^*$ and anticyclical policy generates larger permanent misclassification errors (this corresponds to the case where $\tilde{p}(stab) \geq \tilde{p}(no - stab)$).

If stabilization favors the entrepreneurs, then its effect on the misclassification probability is ambiguous depending on the distribution of the income gains arising from stabilization as well as the distance between the means (the speed of learning) across the two regimes. If the benefits are greatly skewed in favor of the entrepreneurs and the difference in the speeds of learning not very great, stabilization policy may generate a lower permanent rate of misclassification by making it more affordable to stay longer in entrepreneurship. The more extensive sampling may help some high types who are having bad luck "survive" until their luck improves. Of course, in this case the short run mismatches will be even more pronounced.

What is the bottom line of all this analysis? In general, the selection of the sample size involves a trade off between transitory and permanent misallocations. The longer sampling goes on, the greater the persistence of temporary mismatches but the smaller the probability of permanent errors. If g is convex, stabilization policy slows down the speed of learning and worsens the trade off between transitory and permanent mismatches within any activity; and it also worsens the trade off between the same type of mismatch across different activities. If g is concave, stabilization policy improves economic efficiency.

In order to offer a welfare assessment of stabilization policy we need to know its relative effects on the payoffs to the various activities and also possess a social welfare function that evaluates both the short and the long term mismatches in the various activities. In the absence of this information, one may want to consider arbitrary interesting cases. One such case arises when, due to -perhaps non linear- time dependent costs of career switching, a permanent occupational decision has to be made within a fixed period whose end point, N^*, lies ahead of the optimal switch time. The transitional inefficiencies are the same under both regimes (the length of sampling is the same). But the long term inefficiencies are different depending on the speed of learning. Another case involves equiprobable priors and symmetric loss functions. Under these conditions, both the permanent and transitory inefficiencies are different depending on the exercise of successful macroeconomic stabilization.

Possible extensions

The model can be extended in several directions. One could involve changing the decision problem itself. For instance, one might consider situations in which an individual can chose not only which activity to sample in but also the intensity of effort to devote to that activity (or to a portfolio of activities). The degree of learning could then be affected by the intensity of the effort. Another route might involve the introduction of bias in the exercise of stabilization policy. In general, policymakers may be much more concerned about downturns than upturns. If g is convex, selective stabilization policy can deprive an economy from the most informative outcomes without compensating at the high end. It can be shown that in such a case, a sufficiently asymmetric, anticyclical policy can reduce informational efficiency even in cases where g is linear. Finally, serial correlation in the state of the business cycle could be introduced. A persistent business cycle would add one more state variable to the system and would affect the persistence of mismatches and probability of misclassification.

2. Empirical implications

The function of cyclical relative performance, g, is the building block of our model. The sign of its first derivative determines whether recessions carry superior information for type discrimination (they do if g is decreasing in the favorableness of the business cycle). The sign of its second derivative determines the net -over the cycle- effect of stabilization policy on informational efficiency (the effect is negative if g is convex). Unfortunately, as types are not directly observable, g is not observable either. In this section we suggest possible ways for inferring its shape by relying on some important predictions of the theory: one concerns the properties of the empirical distribution of performance of firms (or individuals) in any activity; another concerns the implied pattern of business dissolutions.

I. Implications for the empirical distribution of performance

The properties of the function of relative performance can be deduced from the cyclical behavior of the empirical distribution of performance in any activity.

Proposition 7 *Let* $X_n = (x_1, x_2, .., x_n)$ *and* $Y_m = (y_1, y_2, .., y_m)$ *be observations of performance of two individuals of different type in the same activity. Let* μ_h *and* μ_l *be the mean and* σ_h *and* σ_l *the variance of* X_n *and* Y_m *respectively. Consider the union of the two samples,* $Q_M = X_n \cup Y_m$, $M = n + m$. *Let* $q_i \in Q_M$ *where* q_i *is not known whether it came from* X_n *or* Y_m *and let* Θ *denote the probability distribution of* q.

a) $E\ q = \mu = (n/M)\mu_h + (m/M)\mu_l$, $Var(q) = \sigma = (n/M)\sigma_h + (m/M)\sigma_l + (\mu_h - \mu_l)^2/2$.

b) Let $q(r)$ *be the* r*-th quantile of* q_i. *Furthermore, assume that* q_i *is normal. Then*

$d \mid q(r) - q(1/2) \mid / d \mid \mu_h - \mu_l \mid > 0$.

Proposition 7 suggests the following way of inferring the shape of g. If g is countercylical ($g'(s) < 0$) then the variance of the empirical probability distribution within any activity is increasing in the severity of the recession (recall that the sample means are further apart during recessions when g is decreasing in s). Moreover, the rank statistics of this distribution ought to grow further apart as times worsened. It is easy to show that the curvature of g can be inferred from the sign of $d^2\sigma/ds^2$ and $d(\overline{q} - \mu)^2/ds^2$(or from the behavior of the distance between any pair of percentiles).

I. Implications for firm exit

The model can be usefully applied to study the patterns of firm exit. According to the theory developed here, a business closure is the result of accumulated information indicating that the chosen activity does not represent the best match[10]. The simplest way to relate the shape of the function of firm closures to the shape of the function of relative performance is by modifying the assumption that the business cycle shock is serially uncorrelated. It can be shown that with positive serial correlation in the state of the business cycle, activity switches ("business failures") are more likely to occur during bad times. And moreover, that the shape of the function of cyclical business closures will match the shape of the function of relative performance[11]. Before concluding this section it is important to make one final observation. The model carries implications for the relationship between volatility and short and long term performance. Let us denote by SR, LR and VSR some measure of short term performance, long term performance and the variance of short term performance respectively (the growth rate or the level of output might be such a measure). If $g(s)$ is decreasing then there is a negative relationship between SR and VSR (because of Proposition 7). Moreover, if $g(s)$ is convex then there exists a positive relationship between VSR and LR. Economies which experience greater short term cyclical volatility ought to perform better in the long run. These predictions can also be used as a means of inferring the shape of $g(s)$. Finally, note that unlike the rational expectations imperfect information models of the business cycle in which aggregate fluctuations are a source of greater confusion, in our model aggregate fluctuations reduce noise.

3. Empirical results

We now turn to the empirical evidence. One should be aware of two possible problems: one is that the variance of the idiosyncratic shock may be an increasing function of the severity of downturns; and the other is that the curvature of the function of relative performance may not be invariant to monotonic transformations of the macroeconomic adversity variable (for instance, the results may differ depending on whether one uses levels, log-levels, growth rates etc.). The latter problem is present in all non-linear models.

We pursue two empirical strategies. The first one is based on the study of the cyclical behavior of business failures in the US[12]

The data represent the number of failures per 10,000 concerns (the failure rate). They are monthly, seasonally adjusted, extend from 1947:01 to 1978:12 and were compiled by Dun and Bradstreet (source: Business Statistics of the Survey of Current Business). Table 1 reports the results from a regression of the failure rate on the index of industrial production (both series have been detrended using the HP filter). A quadratic term is included in order to capture the curvature of the g function. The results -which are very robust to changes in the specification- indicate that the failure rate is a decreasing function of the favorableness of the business cycle[13] ($g' < 0$), and moreover, that $g" > 0$.

The second empirical strategy is based on proposition 7. We calculate the empirical distribution function of the growth rate in annual sales between 1970 and 1990 for the companies contained in the Compustat data base.

The annual sales growth variable used is a symmetric, bounded variable that is a monotonic transformation of the standard growth rate (so that increases and decreases in sales are treated symmetrically). It was constructed as $g_t = (x_t - x_{t-1})/.5(x_t + x_{t-1})$, where x_t is real sales in period t for the firm under consideration.

Table 3 reports the results from regressions of the -annual- differences between various percentiles of the empirical distribution of sales on the cross-industry average growth rate of sales. Two samples were employed. One consisted of all stock exchange listed firms; and the other of its subset of manufacturing companies. The t- statistics ought to be interpreted with some caution and only as indicative because the properties of the estimated coefficients are not known. Nonetheless, it is worthwhile noting that the results uniformly indicate that the difference in relative performance across the percentiles is countercyclical. Decreases in average "aggregate" performance are associated with a spread out of the empirical distribution. Similar findings apply to the relationship between the mean (or median) and the standard deviation of sales. Note, though, that no nonlinear pattern is detected when the regressions are augmented to include higher order polynomial terms of the average growth rate (that is, the assumption that g is linear cannot be rejected).

In order for this piece of evidence to indicate that recessions are more informative, it must also be the case that, on average, the same companies remain at the top (or at the bottom) of the distribution of performance during different recessions or expansions. The examination of the frequency of switches between above median and below median performance shows that this is indeed the case (the same pattern also obtains within sub industries in the sense that some firms systematically outperform other firms).

The preceding analysis has been implicitly based on the assumption that a company's reference group is the entire population of firms rather than the own industry firms. It is not obvious what the relevant group (the right pond) is in a multisector economy. For instance, the goodness of performance could be judged against the performance of the other firms in the same sector rather than in the economy at large. Regressions similar to those reported in Table 3 were run within individual two and four digit industries; and by also pooling across industries but using for each industry the corresponding industry rank statistics (and mean). Unlike the first set of regressions (Table 3), the new regressions, with the exception of the chemicals industry, did not produce any evidence that the distributions varied systematically over the business cycle (see Table 4). One must note, however, that most of the intra-industry analysis is problematic because of the very small number of observations.

Note also that the Compustat data set is far from ideal for our purposes because it includes only large successful firms. As a result, the cross variation of sales within the sample of companies over the business cycle may be significantly understated. Another problem arises from the fact that the sample leaves out the bottom of the distribution, that is, firms that failed. Both of these shortcomings of the sample tend to bias our results against showing a significant increase in the dispersion of the distribution of sales during recessions (as measured by the variance or the distance between the various quantiles); and against finding a convex pattern in the difference of performance between the top and the bottom of the distribution.

Conclusions

An economy can achieve a more efficient allocation of resources when economic agents find out early and accurately whether they have made the right occupational or entrepreneurial choices. One important way of learning about who is what and where one belongs is by sampling one's performance in a selected activity and drawing the appropriate inferences. Prolonged sampling increases the ability to discriminate and avoid long term mismatches. But this comes at the expense of more persistent transitory mismatches. If the individuals who have made the wrong decisions suffer relatively more when the times are bad, then recessions may carry superior discrimination information. If such effects are pronounced, then stabilization policy may worsen the trade off between short and long term misallocation.

The empirical work carried out provides some support to the notion that aggregate "bad" times possess greater revelation properties than "good" times and hence recessions can play an important role in separating the wheat from the chaff. The evidence on the informational implications of stabilization policy is less clear.

There now exist several theories in the literature linking the short to the long run. Which of the suggested mechanisms -if any- is the most relevant one from an empirical point of view remains an open issue. There is little doubt, however, that the informational-learning considerations emphasized here are at the heart of many important human decisions.

References

Anderson, T. ,1958, An introduction to multivariate statistical analysis, Wiley, New York.

Berger, J., 1985 Statistical Decision Theory and Bayesian Analysis, Berlin: Springer-Verlag.

Schumpeter, J., 1991, Essays, edited by R. V. Clemence, Oxford: Transaction Publishers.

Wald, A., 1947, "Sequential decision," New York: J. Wiley.

APPENDIX

Proposition 1: There exists a $\tilde{p}(X_t)$:

$E_t[Y_e(p_t, s_t, u_{et}) + \beta V_e(p_{t+1}, s_{t+1}, u_{et+1})] \geq E_t[V_w(p_t, s_t, u_{wt})\ for\ p_t \geq \tilde{p}$

$E_t[Y_e(p_t, s_t, u_{et}) + \beta V_e(p_{t+1}, s_{t+1}, u_{et+1})] < E_t[V_w(p_t, s_t, u_{wt})\ for\ p_t < \tilde{p}$

Proof: The proof follows immediately from the fact that the expected lifetime income of one who chooses to prolong his stay in entrepreneurship by one more period (the LHS of the inequality above) is a strictly increasing function of p; the expected lifetime income of one who switches to workmanship (the RHS of the inequality) is a strictly decreasing function of p; and assumption 1.

Proposition 4: For any sample size t, the average value of LF_t is decreasing in $\mid \mu_h^* - \mu_l^* \mid$ when the true type is h (increasing when true type is l).

Proof: Let $P(X_t|\mu)$ be the likelihood function of μ associated with the sample $X_t = \{x_1, x_2, \ldots, x_t\}$ of normal variables. $P(X_t|\mu)$ is maximized when $S(\mu, X, t) = \Sigma_1^t(\mathrm{x}_i\text{-}\mu)^2$ is minimized. S is a strictly convex function of μ with a minimum at $\mu = \mu_h^*$ when true type is h. Consequently, $d[P(\mu_l^*)/P(\mu_h^*)]/d \mid \mu_h^* - \mu_l^* \mid < 0$ and thus $LF_t = p(X|\mu_l^*)/p(X|\mu_h^*)$ is a decreasing function of $\mid \mu_h^* - \mu_l^* \mid$ when true type is h. A similar argument can be used to establish that LF_t is an increasing function of $\mid \mu_h^* - \mu_l^* \mid$ when the true type is low.

Proposition 5: The approximate expected stopping time is given by

$$E(N) = \frac{E \ln R}{E(z)}$$

Proof: Recall that our subject of study is a low type individual who is sampling in entrepreneurship. Subsequently, for such an agent $E(z) = -(1/2\sigma^2)(\mu_h^* - \mu_l^*)^2 < 0$. To establish that $lnR < 0$ note that the optimal rule for initial selection implies that entrepreneurship is the activity of choice if $p(h) > A_1/(1 + A_1)$. Using the fact that $p(l) + p(h) = 1$ and $A_1 > A$, the initial selection rule can be rewritten as $1 > A_1[p(l)/p(h)] > A[p(l)/p(h)] \equiv R$.

Proposition 6: Let g be a convex function of the business cycle. The probability of permanent misclassification is increasing in the degree of stabilization if the exercise of stabilization policy does not favor the entrepreneurs.

Proof: $(\xi_t - \mu_h^*) = -(\mu_h^* - \mu_l^*)/2 - [(\sigma^2/t)/(\mu_h^* - \mu_l^*)]lnR$. F is increasing. We have shown that $lnR > 0$ (otherwise the individual would not be in entrepreneurship). Whether $\xi_t - \mu_h^*$ is decreasing in $\mu_h^* - \mu_l^*$ or not depends on $B = sign\ \{dlnR/d[\mu_h^* - \mu_l^*]\}$. If stabilization helps workers at least as much as it helps entrepreneurs then $B > 0$. Consequently, $\xi_t - \mu_h^*$ is decreasing in $\mu_h^* - \mu_l^*$ and anticyclical policy generates larger permanent misclassification errors (this corresponds to the case where $\tilde{p}(stab) \geq \tilde{p}(no - stab)$). If $B < 0$, the effect is ambiguous.

Proposition 7: Let $X_n = (x_1, x_2, .., x_n)$ and $Y_m = (y_1, y_2, .., y_m)$ be observations of performance of two individuals of different type in the same activity. Let μ_h and μ_l be the mean and σ_h and σ_l the variance of X_n and Y_m respectively. Consider the union of the two samples, $Q_M = X_n \cup Y_m$, $M = n + m$. Let $q_i \in Q_M$ where q_i is not known whether it came from X_n or $Y_m$and let Θ denote the probability distribution of q.
a) $E\ q = \mu = (n/M)\mu_h + (m/M)\mu_l$, $Var(q) = \sigma = (n/M)\sigma_h + (m/M)\sigma_l + (\mu_h - \mu_l)^2/2$.
b) Let $q(r)$ be the r-th quantile of q_i. Furthermore, assume that q_i is normal. Then $d \mid q(r) - q(1/2) \mid /d \mid \mu_h - \mu_l \mid > 0$.

Proof: a) The derivation of $E(q)$ is straightforward. The variance of q is

$$\sigma = [\Sigma_1^M (q_i - \mu)^2]/M = [\Sigma_1^n (x_i - \mu)^2 + \Sigma_1^m (y_i - \mu)^2]/M =$$

$$[\Sigma_1^n (x_i - \mu_h + (m/M)(\mu_h - \mu_l))^2 + \Sigma_1^m (y_i - \mu_l + (n/M)(\mu_l - \mu_h))^2]/M =$$

$$(n/M)\sigma_h + (m/M)\sigma_l + (n/M)(m/M)(\mu_h - \mu_l)^2/2.$$

b) Let $prob(q < \overline{q}) = prob[\sqrt{M}(q - \mu)/\sigma < \sqrt{M}(\overline{q} - \mu)/\sigma) = \Theta(\sqrt{M}(\overline{q} - \mu)/\sigma) = \Theta(\overline{q})$. Consider an increase in $\mu_h - \mu_l$ that leaves μ unchanged. Θ is increasing. Part (a) implies that σ increases and consequently, for a given $\Theta(\overline{q})$, $\overline{q} - \mu$ must increase too.

Table 1: Business Failures and the Business Cycle, 1947-83

	Constant	BF_{t-1}	IP_t	$IP_t{*}IP_t$
BF_t	-.010* (.005)	.30** (.04)	-.75** (.14)	6.5** (2.6)
	R^2=.29	s.e.=.08	DW=2.05	Obs: 442

Notes: BF = Failure rate, IP = Index of Industrial Production. BF and IP have been detrended using the HP filter

*Indicates significance at the 5-percent level, **Indicates significance at the 1-percent level

Standard errors in parenthesis

Table 2: Rate of Business Failures

	Mean	Median	Max	Min	Std.Dev.
Failure Rate (Detrended)	-.0036	-.011	.30	-.31	.10

Table 3: Rank Statistics and Average Sales Growth

ALL COMPANIES

	DP90-10	DP90-50	DP50-10	DP75-25	SD
AVG	-.55** (.20)	.-35** (.15)	-.20** (.07)	-.71 (.39)	-.23* (.10)
R^2	.29	.24	.29	.16	.24

MANUFACTURING COMPANIES

	DP90-10	DP90-50	DP75-25	SD
AVG	-.27* (.13)	.-32** (.13)	-.66* (.29)	-.19* (.07)
R^2	.19	.25	.23	.27

DPi-j = Difference between i-th and j-th percentile; AVG(t) = "average" cross sectional growth rate of real sales in year t, SD(t)=standard deviation of cross sectional growth rate of real sales in year t.

Table 4: Rank Statistics and Average Sales Growth: Own Industry Mean, Industry Dummies

	DP90-10	DP75-25
AVG	-.05 (.12)	.-06 (.22)
R^2	.69	.45

Notes

1. Whether this occupational choice will be irreversible or not depends on whether there exist costs to switching occupation after some period, on whether sampling also takes place in the new activity and so on.

2. According to the popular saying "when the going gets tough the tough get going". In our story, toughness corresponds to absolute advantage.

3. It must be stressed that our argument does not require taking a position on whether it is expansions or contractions that carry greater information. That is, it may well be the case that differences in performance are greater during good times. What is important is that there is a difference.

4. Allowing for multiple switches may be useful for accounting for -perhaps periodic- career changes over the business cycle. However, it makes the analysis more cumbersome without bringing in any additional insights to the analysis of the informational content of the business cycle. We have opted for a single switch to make the distinction between transitory and permanent mismatches sharp.

5. We are describing the behavior of an agent who has not already made a career switch.

6. Note that the initial period priors are used in forming the current period's posterior belief. We could have alternatively used last period's posterior as the current prior. While the latter specification would lead to faster learning for a given sequence of s_t's, it would not change any of the results concerning the effects of stabilization policy on the speed of learning and the probability of misclassification (which is the subject of this paper). The likelihood function specification was adopted because it is somewhat simpler.

7. A similar rule describes the behaviour of a worker.

8. Even in the absence of normality the posterior mean will be approximately normally

distributed because of the central limit theorem. However, it would not necessarily be the optimal estimator of type unless the loss function were quadratic.

9. Stabilization policy may take the form of aggregate demand stabilization, product or input price controls and so on.

10. We do not claim that all business dissolutions in the real world are associated with learning. Many other reasons exist. For instance, a firm may have been established to take advantage of a temporary opportunity and ceases its operation once it has fulfilled its objectives.

11. A possible practical difficulty in inferring $g(s)$ from the pattern of business failures arises from the fact that the shape of g may depend on the measure of the business cycle (s) used. For instance, $g(s)$ may be convex in s but if one uses another business cycle variable, $\lambda = \psi(s)$,then $g(\lambda)$ may not be.

12. A failure is defined as "a concern that is involved in a court proceeding or a voluntary action that is likely to end in a loss to creditors." The failures data exclude railroads, banks, financial companies, holding companies real estate and insurance brokers, amusement enterprises, shopping agents tourist companies and transportation terminals.

13. Similar results obtain when the cycle is defined according to the $NBER$ chronology.