



HAL
open science

Free entry and business cycles under the influence of animal spirits

Rodolphe dos Santos Ferreira, Frédéric Dufourt

► **To cite this version:**

Rodolphe dos Santos Ferreira, Frédéric Dufourt. Free entry and business cycles under the influence of animal spirits. *Journal of Monetary Economics*, 2006, 53 (2), pp.311-328. 10.1016/j.jmoneco.2004.12.003 . halshs-00789030

HAL Id: halshs-00789030

<https://shs.hal.science/halshs-00789030>

Submitted on 15 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Free entry and business cycles under the influence of animal spirits*

Rodolphe Dos Santos Ferreira[†] Frédéric Dufourt[‡]

Revised version - August 2004

First version - July 2002

Abstract

We provide a business cycle model in which endogenous markup fluctuations are the main driving force. These fluctuations occur due to some form of ‘animal spirits’, impelling firms in their entry-exit decisions within each sector. By contrast to existing models of the business cycle emphasizing the role of animal spirits, we do not rely on the sink property of the equilibrium to generate indeterminacy. Hence, while our model does pretty well in accounting for the main features of US business cycles, it avoids several criticisms addressed to these former models, concerning either their dependence upon strongly increasing returns and too high markups, or their implication of countercyclical movements of consumption.

JEL classification number: D43, E32, L1

Keywords: Business cycles, Animal spirits, Imperfect competition.

*Forthcoming in the *Journal of Monetary Economics*, 2006, Vol. 53(2). We thank an anonymous referee for valuable comments and suggestions.

[†]BETA-Theme, University Louis Pasteur and Institut Universitaire de France. *Corresponding author*. Address: BETA -61 avenue de la Forêt Noire, F-67085 Strasbourg, France, Tel: 33-3-90 242073. Fax: 33-3-90 242071. E-mail: rdsf@cournot.u-strasbg.fr

[‡]BETA-Theme, University Louis Pasteur, Strasbourg. E-mail: dufourt@cournot.u-strasbg.fr

1 Introduction

We propose a business cycle model in which the main source of fluctuations consists in “animal spirits”, as they influence firms’ willingness to enter and compete within existing markets. Although, like the previous literature on endogenous fluctuations, we focus on “animal spirits” as an important factor of the business cycle,¹ we do not rely on the sink property of the dynamic equilibrium to generate indeterminacy. We use instead previous microeconomic theoretical results showing that, under private increasing returns to scale, free entry and perfect market contestability are compatible with an indeterminate number of active firms at equilibrium. To be precise, we mean that different levels of activity, corresponding to different numbers of producing firms within a larger set of potential competitors, are sustainable as Nash equilibria.

The usual approach to free entry is based upon the view that positive profits of incumbent firms are always a sufficient incentive for potential entrants to become active, so that equilibrium in a contestable market is possible only under the zero profit condition. This condition is then used to endogenously determine the number of active firms at equilibrium. However, the standard concept of Nash equilibrium, applying to both busy and idle firms, only requires that there be no admissible way for the latter to attain a positive profit, given the equilibrium strategies of the former. While this is true under different forms of competition, the Cournot equilibrium with free entry (Novshek, 1980)² offers a straightforward illustration of this idea. If profits are positive and if the optimal individual scale is negligible with respect to market size, nothing prevents a potential entrant to produce at a near optimal scale and sell at an almost unchanged price. Hence, equilibrium profits are necessarily close to zero. However, if the optimal individual scale is non-negligible, as implied by the existence of internal economies of scale, then not only can active firms’ profits be significantly different from zero, but the equilibrium number of these firms may also be largely indeterminate (d’Aspremont, Dos Santos Ferreira and Gérard-Varet, 2000).

As far as we know, no attempts have been made to explore the implica-

¹A non-exhaustive list of papers on this issue includes Benhabib and Farmer (1994), Farmer and Guo (1994), Galí (1994), Benhabib and Farmer (1996), Schmitt-Grohé (1997), Perli (1998), Wen (1998), Schmitt-Grohé (2000) and Weder (2000). See the very complete survey by Benhabib and Farmer (1999) for other important references.

²See also Knieps and Vogelsang (1982), and Brock and Scheinkman (1983).

tions of these microeconomic results at the macroeconomic level, in particular for business cycle studies.³ The aim of this paper is precisely to explore whether this indeterminacy property may have strong implications for the business cycle. Our main idea is that if the equilibrium number of active firms is indeterminate, the actual number will depend on varying consistent conjecture profiles inducing some firms to produce and at the same time dissuading their potential competitors from doing so. In short, “animal spirits” may influence entry and exit decisions, and so become a driving force of the business cycle. Notice that we are not saying that “animal spirits” are supposed to play the role of exogenous random shocks on the number of active firms, but rather that they may act as a stochastic selection mechanism in the presence of indeterminacy.

A simple empirical implication of our model is that optimistic and pessimistic expectations of firms should be correlated with periods of creation and destruction of business plants, respectively. Figure 1 displays firms expectations about future production together with net business formation in France over the period 1993-2002 (monthly data).⁴ Although the match is far from perfect, due principally to very high frequency movements in the process of net business formation, there is a clear correspondence between the two series. At a monthly frequency the correlation is 0.45, and it raises to 0.53 at a quarterly basis. Figure 2 shows in turn that net business formation is strongly correlated with detrended output (HP-filtered). The correlation is around 0.60, and is similar to that found in Portier (1995) for France during the period 1977-1989, and in Chatterjee and Cooper (1993) for the United States. Of course, it is possible to explain such a relationship by considering that output and business formation simultaneously increase in response to real shocks such as technological innovations.⁵ In this paper, we simply

³In the seminal paper of Chatterjee, Cooper and Ravikumar (1993), strategic complementarities in the entry decisions of consumers-producers may also lead to multiple Pareto-ranked equilibria. However, in that paper, the mechanism generating indeterminacy involves participation costs which are different across agents. This is not in accordance, strictly speaking, with free entry conditions, in the sense given in the microeconomic literature mentioned above.

⁴All the data are provided by INSEE. The expectations series is taken from the monthly survey on manufacturing “expectations about future production”, and the net business formation series is given by “creations less destructions” as recorded by INSEE. The output series is HP-filtered gdp. We thank Guy Laroque and Vladimir Passeron for providing us with these data.

⁵For example, Rotemberg and Woodford (1995) and Portier (1995) show that the num-

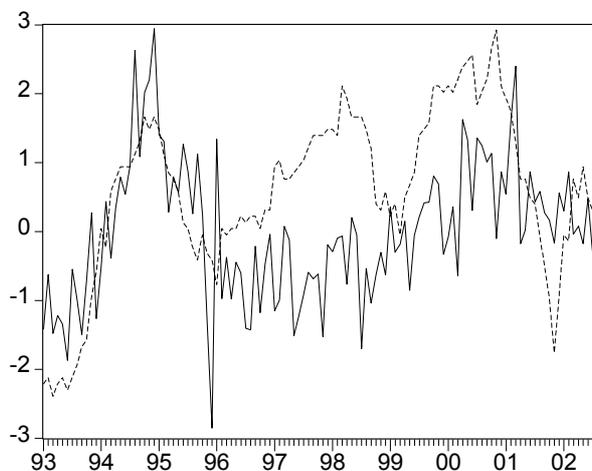


Figure 1: Firms' expectations about future production (dashed line) and net business formation (solid line) in France, 1993-2002.

wish to explore the alternative explanation, which attributes to self-fulfilling expectations of producers the process of entry and exit of new firms within any particular sector.

To illustrate our main idea, we build a simple symmetric model of imperfect competition with a large number of differentiated sectors, within which only a small number of Cournot competitors are active at equilibrium. In line with the microeconomic literature mentioned above, we show that this number may be indeterminate under free entry. Thus, although there is no intrinsic uncertainty, fluctuations in the number of active firms may occur due to endogenous changes in the beliefs of producers with respect to their competitors' behavior. These changes are assumed to be coordinated by reference to some extrinsic stochastic process, with both idiosyncratic and aggregate components. Such a mechanism of entry, taking place in the absence of fundamental uncertainty, allows us to generate endogenous fluctuations in the average markup and, therefore, endogenous fluctuations in the whole economic system. We simulate a dynamic general equilibrium model incorporating such endogenous variations in beliefs, and show that it is able

ber of active firms should vary with output in response to technological and government spending shocks under the zero profit condition.

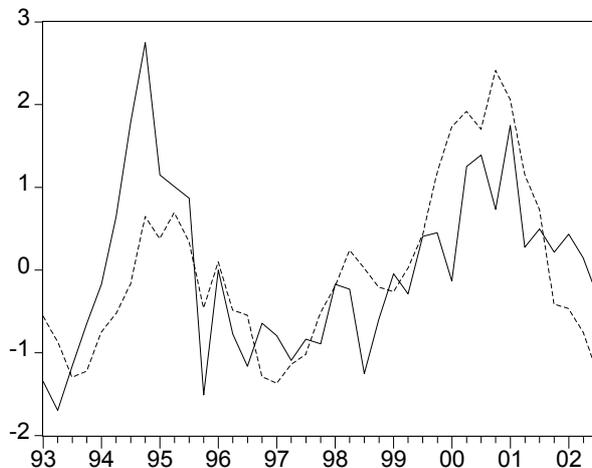


Figure 2: Net business formation (solid line) and HP-filtered output (dashed line), France, 1993-2002.

to generate large fluctuations in real variables. Furthermore, simulation experiments show that the business cycle properties implied by our model are comparable to those measured in the US economy.

The remainder of the paper is organized as follows. Section 2 presents the basic structure of the production sector of our model, and illustrates the indeterminacy result on the equilibrium number of active firms. Section 3 introduces the markup formation process emphasized in section 2 into a dynamic general equilibrium model, and discusses aggregation and coordination issues. Section 4 provides the simulation results and discusses the performance of our model with respect to related models in the literature. Section 5 concludes.

2 Free entry and markup formation

We show in this section that an indeterminate equilibrium number of active firms may prevail in an economy endowed with a quite standard production and market structure. Although the structure we use is chosen so as to make calculations as simple as possible, none of the results here emphasized is specific to that particular structure. In fact, as shown in d'Aspremont, Dos

Santos Ferreira and Gérard-Varet (2000), indeterminacy in the equilibrium number of active firms within perfectly contestable markets may occur under fairly general conditions regarding the market structure and the type of competition which is involved between firms. A necessary condition is obviously that returns to scale be increasing, otherwise a finite number of active firms is incompatible with a free entry equilibrium. However, even a modest degree of increasing returns to scale, resulting from the mere existence of fixed costs, may well be sufficient for indeterminacy.

The economy we consider consists of a large number m of identical sectors, each one producing a homogeneous good, and such that goods stemming from different sectors are imperfectly substitutable from the consumers' viewpoint. In each sector, a virtually large number N of identical firms are involved in Cournot competition within a perfectly contestable output market. Each one of these firms has to decide its output, positive if it chooses to be active, otherwise zero. We begin by examining partial equilibrium conditions in any such sector, leaving their general equilibrium counterpart to the next section.

For simplicity, we assume that individual production takes place under a constant positive real marginal cost c plus a fixed (non sunk) positive cost ϕ , incurred in terms of wasted product. In other words, total real costs to be paid to supply any *positive* amount y of good are given by⁶ $c(y + \phi)$. Furthermore, we assume that the demand addressed to any sector i depends negatively (with unit elasticity) on the price of this sector relative to the general price level, p_i/P , and positively and linearly on an index of sectoral demand, a_i :

$$y_i = \frac{a_i P}{p_i}, \quad (1)$$

where a_i and P are taken as exogenous from the standpoint of any individual sector i .

Given this set of assumptions, we can study the optimal behavior of any firm $j \in \{1, \dots, N\}$ potentially producing in sector i . Any such firm, anticipating aggregate real spending a_i and a vector $y_{i(-j)} = (y_{i1}, \dots, y_{ij-1}, y_{ij+1}, \dots, y_{iN})$ of outputs supplied by its competitors, has to solve the program:

⁶This is the cost structure corresponding to a production function with constant returns to scale relative to variable inputs, when firms have to pay in addition overhead costs ϕ , expressed in terms of their output.

$$\max_{y_{ij} \in [0, \infty)} \left\{ \left| \begin{array}{ll} \left(\frac{a_i}{y_{ij} + \bar{y}_{i(-j)}} - c \right) y_{ij} - c\phi & \text{if } y_{ij} > 0 \\ 0 & \text{if } y_{ij} = 0 \end{array} \right. \right\}, \quad (2)$$

where $\bar{y}_{i(-j)} = \sum_{j' \neq j} y_{ij'}$ is the total amount of output supplied by firm j 's competitors.

According to this program, the optimal behavior of a firm depends on the conjecture it makes on other firms behavior. Naturally, the choices of all other firms are also dependent on the conjectures they make on the former's behavior. These strategic interactions underline the concept of free entry equilibrium, which is nothing else but a qualification of the standard concept of Cournot-Nash equilibrium:

Definition: An *equilibrium* in sector i , given the common expectation of an aggregate sectoral real expenditure a_i , is a vector $(y_{i1}^*, \dots, y_{iN}^*) \in [0, \infty)^N$ such that y_{ij}^* is a solution to producer j 's program (2) under the conjecture $y_{i(-j)}^*$. It is an *equilibrium with free entry* if some firms are active, while at least one other firm optimally chooses not to produce: $\exists j \in \{1, \dots, N\}, y_{ij}^* = 0$, while $\bar{y}_{i(-j)}^* > 0$.

Each one of the N firms within the sector is allowed to rationally choose to be active or inactive, according to its (correct) conjecture of other firms choices. As soon as one firm at least chooses to be inactive (the equilibrium number n of active firms being then smaller than N), we may speak of free entry since there exists a potential entrant which is not hindered by any cost or product differentiation disadvantage relative to its competitors.

We now show how to compute the admissible interval for the equilibrium number of active firms. First, notice that any firm deciding to be active must choose the amount of good it produces according to the first order condition for a positive solution of program (2):

$$\frac{a}{y_j + \bar{y}_{-j}} \left(1 - \frac{y_j}{y_j + \bar{y}_{-j}} \right) = c \quad (3)$$

(where we have dropped for simplicity the subscript referring to the sector). For $\bar{y}_{-j} > 0$, and by concavity of the payoff function in the interval $(0, \infty)$, this condition is sufficient for a global maximum with positive production, provided the corresponding solution is profitable, that is, entails a non-negative profit.

Observe that the left-hand side of (3), with the same equilibrium level of sectoral output $y_j^* + \bar{y}_{-j}^*$ for any j , is decreasing in y_j , so that a positive solution y_n to (2) is unique and identical to all the n active firms, entailing the individual output

$$y_n = \frac{1}{n} \frac{aP}{p_n} \quad (4)$$

and the relative price

$$\frac{p_n}{P} = \frac{n}{n-1} c. \quad (5)$$

Thus, firms set their common price by applying to marginal cost c a markup which, by the assumption of unit demand elasticity, only depends (negatively) upon the number n of active firms. This number is endogenous, and must satisfy, besides the *profitability* condition that any producing firm should make a non-negative profit, a *sustainability* condition for free entry. This condition is weaker than the zero profit condition that is usually introduced in the macroeconomic literature. It merely imposes that any potential entrant is deterred from becoming active because *correctly* realizing that no output would entail a positive profit, given its conjecture of the sectoral demand level a and of the cumulative sectoral output ny_n supplied by its n active competitors.

Within our framework, the *profitability* condition is

$$\left(\frac{a}{ny_n} - c \right) y_n \geq c\phi, \quad (i)$$

whereas the *sustainability* condition is

$$\forall y_j \in [0, \infty), \quad \left(\frac{a}{y_j + ny_n} - c \right) y_j \leq c\phi. \quad (ii)$$

In particular, the left-hand side of the second inequality expresses the profit of a potential entrant, gross of the fixed cost. This profit is maximized at some level of output \hat{y}_n , which is given by

$$\hat{y}_n = \frac{a}{c} \sqrt{\frac{n-1}{n}} \left(1 - \sqrt{\frac{n-1}{n}} \right), \quad (6)$$

so that condition (ii) can be given a more tractable, equivalent expression:

$$\left(\frac{a}{\hat{y}_n + ny_n} - c\right)\hat{y}_n \leq c\phi. \quad (\text{ii}')$$

Condition (ii') simply states that even for the “optimal” level of positive production, a potential entrant can at best exactly cover its production costs, thus being satisfied with inaction.

Conditions (i) and (ii') can now be used to determine the admissible interval for the equilibrium number of active firms. In particular, using (4) and (5) to substitute for y_n and p_n/P , the profitability condition (i) yields

$$n \leq \sqrt{\frac{a}{c\phi}} \equiv \bar{n}. \quad (7)$$

Similarly, using (4) to (6) to substitute for y_n , p_n/P and \hat{y}_n , the sustainability condition (ii') can be expressed as

$$1 - \sqrt{\frac{n-1}{n}} \leq \sqrt{\frac{c\phi}{a}},$$

or

$$n \geq \frac{\bar{n}}{2 - 1/\bar{n}} \equiv \underline{n}. \quad (8)$$

Hence, any integer n in the interval $[\underline{n}, \bar{n}]$ can be the number of active firms at an equilibrium with free entry (there is indeterminacy as soon as the interval contains more than one integer). The zero profit condition, picking up the greatest integer in the interval, thus appears as no more than a particular selection device in presence of indeterminacy.

3 The model

In this section and the following, we study the business cycle properties of a dynamic general equilibrium model which includes markup formation as just characterized. Specifically, we build a standard dynamic general equilibrium model which relies on the market structure described in section 2, and assume that firms enter any sector according to some form of “animal spirits”, selecting a number of active firms in the admissible interval. We think of “animal spirits” as the main mechanism governing entries because, as already stressed, producers’ equilibrium decisions are strongly dependent

on consistent, self-fulfilling *conjectures* about their competitors' behavior. We want further to examine whether such an entry mechanism, associated with extrinsic random shocks on producers' conjectures (and the corresponding markup formation) can provide a model of economic fluctuations whose properties are compatible with those observed in the US economy.

The economy is composed by a huge population (formally, a continuum of unit mass) of identical households maximizing their intertemporal utility, and by a large number m of sectors producing m imperfectly substitutable goods, within which a small number of active firms compete in a Cournot setting. We assume that the utility function of a household is defined, for each period, over the m different goods, with a unit elasticity of intersectoral substitution.⁷ As well known, this specification implies that the total demand addressing each sector i at time t is given by

$$y_{i,t} = \frac{Y_t P_t}{m p_{i,t}}, \quad (9)$$

where $Y_t = m \left[\prod_{i=1}^m (y_{i,t})^{1/m} \right]$ and $P_t = \prod_{i=1}^m (p_{i,t})^{1/m}$ are the appropriate indices of the aggregate output and the price level. Hence, the demand which addresses each sector depends negatively, with unit elasticity, on the relative price $p_{i,t}/P_t$ prevailing in the sector, and positively and linearly on the level $a_{i,t} \equiv Y_t/m$ of sectoral demand, as assumed in the preceding section.

Because of the presence of fixed costs in the production function, there are increasing returns to scale; otherwise, our model (with Cobb-Douglas preferences and technologies, in particular) is quite standard. But this weak form of increasing returns is enough to entail indeterminacy of the equilibrium number of active firms within each sector, so that different levels of activity are sustainable, even though output markets are all perfectly contestable.

3.1 Households

The representative household is endowed with an initial amount of capital K_0 . During each period t , $t = 0, \dots, \infty$, it rents its capital stock K_t to the representative firm at the real interest rate r_t , supplies an amount of work H_t at the real wage rate w_t , perceives as a shareholder the real profits made

⁷As mentioned earlier, a unit elasticity of intersectoral substitution is one of the assumptions made only to simplify calculations. None of our results on the indeterminate equilibrium number of firms is dependent on this particular assumption.

by the representative firm Π_t , and consumes a volume C_t of the final goods. The intertemporal budget constraint is

$$K_{t+1} = (1 - \delta)K_t + r_t K_t + w_t H_t + \Pi_t - C_t, \quad (10)$$

where δ is the rate of capital depreciation. The program of the representative household is then to maximize the expected utility

$$E_0 \left\{ \sum_{t=0}^{\infty} \beta^t U(C_t, H_t) \right\}$$

with respect to $\{C_t, H_t, K_{t+1}\}_{t=0, \dots, \infty}$, given the budget constraint (10) and the instantaneous utility function (see Hansen, 1985)

$$U(C_t, H_t) = \frac{1}{1 - \sigma} C_t^{1 - \sigma} - \frac{B}{1 + \chi} H_t^{1 + \chi},$$

where $\sigma > 0$ is the inverse of the elasticity of intertemporal substitution, and $\chi > 0$ is the inverse of the labor supply elasticity. The optimality conditions can be written as

$$BH_t^\chi C_t^\sigma = w_t \quad (11)$$

and

$$C_t^{-\sigma} = \beta E_t \left\{ (1 - \delta + r_{t+1}) C_{t+1}^{-\sigma} \right\}, \quad (12)$$

which are the traditional consumption-leisure and consumption-saving trade-off conditions for the household.

3.2 Firms

The profit maximization program of the representative firm j in the i -th sector can be conveniently described as a two-stage procedure: In the first stage it chooses, for any given level of production $y_{ij,t}$, the optimal combination of capital $k_{ij,t}$ and labor $h_{ij,t}$ which minimizes its production costs $r_t k_{ij,t} + w_t h_{ij,t}$, subject to the production function

$$y_{ij,t} = (k_{ij,t})^\alpha (h_{ij,t})^{1 - \alpha} - \phi, \quad (13)$$

where ϕ is a fixed cost (in terms of output). The optimal levels of capital and labor are given by

$$k_{ij,t} = \left(\frac{\alpha}{1-\alpha} \right)^{1-\alpha} \left(\frac{w_t}{r_t} \right)^{1-\alpha} (y_{ij,t} + \phi) \quad (14)$$

and

$$h_{ij,t} = \left(\frac{\alpha}{1-\alpha} \right)^{-\alpha} \left(\frac{w_t}{r_t} \right)^{-\alpha} (y_{ij,t} + \phi). \quad (15)$$

We can deduce from these equations the corresponding level of the real production cost (for $y_{ij,t} > 0$):

$$r_t k_{ij,t} + w_t h_{ij,t} = A r_t^\alpha w_t^{1-\alpha} (y_{ij,t} + \phi) \equiv c(r_t, w_t) (y_{ij,t} + \phi), \quad (16)$$

where $A = \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)}$ is a constant related to technology conditions. Notice that the cost structure which is defined by (16) is the one we have adopted in section 2. The real marginal cost is constant at the firm level (w_t and r_t are taken as given from the individual point of view), but the fixed cost in the production function implies that increasing returns to scale prevail in the (weak) form of a decreasing real average cost.

In the second stage of its maximization program, each firm chooses the price it will charge and the corresponding level of production. It is the description of this second stage that we have undertaken in the preceding section. The optimal pricing decision of an active firm is to apply a markup over the marginal cost, leading in a sectoral symmetric equilibrium, by equation (5), to the relative price

$$\frac{p_{i,t}}{P_t} = \mu_{i,t} c(r_t, w_t). \quad (17)$$

The markup depends exclusively, because of constant unit demand elasticity, on the actual number of firms operating within the sector:

$$\mu_{i,t} = \frac{n_{i,t}}{n_{i,t} - 1}, \quad (18)$$

where $n_{i,t}$ takes values in the interval $[\underline{n}, \bar{n}]$.

3.3 Aggregation

We have seen that there are in general multiple Nash equilibria in each sector, with different numbers of producing firms (belonging to the interval $[\underline{n}, \bar{n}]$)

and corresponding different relative prices and levels of activity. Denoting by $\lceil \underline{n} \rceil$ the least integer not smaller than \underline{n} and by $\lfloor \bar{n} \rfloor$ the greatest integer not larger than \bar{n} , there actually exists $S \equiv \lfloor \bar{n} \rfloor - \lceil \underline{n} \rceil + 1$ different possible states for any given sector, such that we may associate with any state s ($s = 1, \dots, S$) a specific admissible number $n_s \equiv \lceil \underline{n} \rceil + s - 1$ of active firms, and a corresponding markup $\mu_s \equiv n_s / (n_s - 1)$.

Aggregation can be easily achieved as follows. Denoting by $f_{s,t}$ the proportion of sectors in the economy which are in state s at date t , we can determine the average number of active firms, $n_t \equiv \sum_{s=1}^S f_{s,t} n_s$, and the average markup μ_t , defined as the weighted geometric mean of the S different admissible markups, using equation (18): $\ln \mu_t = \sum_{s=1}^S f_{s,t} \ln (n_s / (n_s - 1))$. Aggregate output can be correspondingly defined, being a weighted geometric mean of the S different admissible sectoral production levels.

Clearly, the evolution both of the average number of active firms and of the average markup, as well as that of all the aggregate variables, *crucially depends on the evolution of the vector* $f_t \equiv (f_{1,t}, \dots, f_{S,t})$, itself determined by the process coordinating producers' conjectures.

3.4 Animal spirits

Because of the existence of multiple Nash equilibria within each sector, there exists a coordination problem between firms, which cannot be solved by referring solely to current objective economic conditions or “fundamentals” of the system. Instead, ‘animal spirits’ of competitors may influence entry and exit decisions, and so become a driving force of the business cycle.

Following a common practice in the literature, we shall consequently assume that firms tackle this coordination problem by referring to some extrinsic stochastic process, which may refer to both idiosyncratic and aggregate components (relevant information is largely sector specific, but the general economic situation significantly influences the way such information is evaluated and interpreted).

Specifically, we assume that at each date t all firms in each sector i receive a signal $s_{i,t} \in \{1, \dots, S\}$ suggesting the state which is most likely to be realized at this date in this specific sector. However, this signal $s_{i,t}$ is assumed to be noisy, so that it can only trigger coordination upon the corresponding state if its intensity is above some threshold, say ρ in a $[0, 1]$ scale. Otherwise, if the signal appears too noisy, firms will prefer to coordinate upon the state

$s_{i,t-1}^*$ that has been observed in the previous period.⁸

We assume that $s_{i,t}$ is the outcome of an *independent non-stationary* Markov process, characterized by a transition matrix Π_ω , common to all sectors and itself generated at the economy level by some *stationary* stochastic process with sample space Ω .⁹ If signal intensities are distributed over the interval $[0, 1]$ uniformly as well as independently across the different sectors, we obtain by the law of large numbers the following evolution of the column vector $\mathbf{f}_t = (f_{1,t}, \dots, f_{S,t})'$:

$$\mathbf{f}_t = [\rho I + (1 - \rho) \Pi_\omega] \mathbf{f}_{t-1}, \quad (19)$$

where Π_ω is an $(S \times S)$ column-stochastic matrix (the s -th element of the s' -th column of Π_ω represents the probability for any firm in a sector of receiving the signal s at period t , conditional on being in state $s' \in \{1, \dots, S\}$ at $t - 1$, when the economy as a whole is in state $\omega \in \Omega$ at t). By pre-multiplying the two sides of this equation by the row vector $\mathbf{n} = (n_1, \dots, n_S)$, we obtain the corresponding evolution of the average number of active firms

$$n_t = \rho n_{t-1} + (1 - \rho) \mathbf{n} \Pi_\omega \mathbf{f}_{t-1}. \quad (20)$$

Clearly, as long as the transition matrix Π_ω stems from a non-degenerate stochastic process, both the average number of active firms and the average markup will be continuous random variables over the supports $[n_1, n_S]$ and $[n_S / (n_S - 1), n_1 / (n_1 - 1)]$, respectively.¹⁰ Also, the time properties of n_t and μ_t will crucially depend on the properties of the stochastic process Π_ω .

Examples

To illustrate our argument, we can think of different sensible coordination processes implying distinct stochastic processes for the average number of active firms.

⁸This assumption, which seems *a priori* natural, allows us to account for the persistence in firms' expectations clearly apparent in the data (see Figure 1).

⁹We may for instance refer to the *double chain Markov model*, where two Markov chains are superposed, one controlling the transition process between the transition matrices of the other (see Berchtold, 1999).

¹⁰With a constant (and regular) transition matrix $\rho I + (1 - \rho) \Pi$, the frequency vector \mathbf{f}_t would converge to a fixed point \mathbf{f} , so that aggregate fluctuations would eventually fade away.

As a first example, we can assume that there are two states summing up the overall opinion about the general tendency of the economy: “good times” and “bad times”. For instance, “good times” may uniformly encourage the entry of some potential producer, provided profitability is preserved, and “bad times” may prompt some heretofore active firm to cease production, provided sustainability is preserved. In other words, all sectors receive the signal $s_{i,t} = \min \{s_{i,t-1} + 1, S\}$ in “good times”, and $s_{i,t} = \max \{s_{i,t-1} - 1, 1\}$ in “bad times”. Of course, the degree of reliability of this signal is still evaluated on the interval $[0, 1]$, triggering a state transition only if it is larger than the threshold ρ . Under these assumptions, it is straightforward to show that the number of active firms in each sector follows a random walk with two delaying barriers n_1 and n_S (see for instance Spitzer, 1976). As for the average number of active firms, notice that we obtain in this case for the product $\mathbf{n}\Pi_\omega$ in equation (20): $\mathbf{n}\Pi_g = \mathbf{n} + [1 \dots 1 \ 0]$ if $\omega = g$ (“good times”) and $\mathbf{n}\Pi_b = \mathbf{n} - [0 \ 1 \dots 1]$ if $\omega = b$ (“bad times”). As a consequence, the aggregate process described by this equation takes the form:

$$n_t = n_{t-1} + (1 - \rho)(1 - f_{S,t-1}) \text{ if } \omega = g \quad (21)$$

$$n_t = n_{t-1} - (1 - \rho)(1 - f_{1,t-1}) \text{ if } \omega = b. \quad (22)$$

The evolution of the average number of active firms thus looks like a standard random walk with two shocks. However this is partly misleading since, because of the delaying barriers, the aggregate process depends upon past microeconomic information on the proportions $f_{1,t-1}$ and $f_{S,t-1}$ of sectors that were in the two extreme states 1 and S at date $t - 1$.¹¹

As a second example, we can refer to a case in which present aggregate information is sufficient to characterize the stochastic process for the average number of active firms. In particular, we can think of a situation in which all firms in the economy receive the same signal $s_{i,t}$ independently of the state prevailing previously in their respective sectors, but with a degree of confidence that remains idiosyncratic. In this case, all the S columns of Π_ω are equal to the same (degenerate) column probability vector $\boldsymbol{\pi}_\omega$ belonging to the canonical basis of \mathbb{R}^S , so that equation (20) becomes

$$n_t = \rho n_{t-1} + (1 - \rho) n_s, \quad (23)$$

¹¹This is therefore a case in which correct aggregation cannot be done without reference to the position of microeconomic units, see for instance Caballero (1992).

where $n_s \equiv \mathbf{n}\boldsymbol{\pi}_\omega \in \{n_1, n_S\}$. We obtain in this case a dynamic stochastic equation in the *aggregate* variable n_t , which is a standard AR(1) process with a finite number of shocks equal to $\#\Omega \leq S$. A simple instance of this case, that we use in our simulations below, is the discrete uniform distribution over the $\#\Omega = S$ possible states.¹²

4 Model properties

In this final section, we proceed to a quantitative evaluation of the model in the spirit of the Real Business Cycle literature. In particular, we ask whether this simple mechanism of entries and exits driven by self-fulfilling changes in firms' expectations can generate business cycles that are relatively close to those observed in the US economy. Testing the business cycle properties of this model is important since, as discussed below, traditional sunspot-driven models based on the sink nature of the steady state have encountered difficulties in explaining some standard features of observed business cycles. Since our model does not rely on that kind of dynamic indeterminacy, one can expect its business cycle properties to be quite different from those in the traditional literature on endogenous fluctuations.

For simplicity, and to allow easier comparison with standard RBC models, we will from now on consider a coordination scheme inducing, as in (23), a dynamic stochastic equation in the *aggregate* variable n_t , of the form:

$$n_t = \rho n_{t-1} + (1 - \rho)u_t, \quad (24)$$

where the parameter $\rho \in (0, 1)$ controls the degree of persistence, and where u_t is a discrete random variable uniformly distributed over the interval $\{[\underline{n}], [\bar{n}]\}$. Notice that there is a deterministic mean value \bar{n} around which the dispersion of the number of active firms remains constant. Indeed, specification (24) implies

¹²In fact, the assumption that all firms in the economy receive the same signal is only auxiliary. The important requirement to obtain a dynamic stochastic equation in the *aggregate* variable n_t (independently of any microeconomic information) is that the probability distribution over the S states be independent of the previous state (or equivalently that all the S columns of Π_ω be equal, for any $\omega \in \Omega$, to the same column probability vector $\boldsymbol{\pi}_\omega$). For instance, if $\boldsymbol{\pi}_\omega = (1 - \omega, 0, \dots, 0, \omega)'$, where ω is uniformly distributed over $[0, 1]$, the variable $n_\omega \equiv \mathbf{n}\boldsymbol{\pi}_\omega$ is a *continuous* uniform random variable.

$$E(n_t) = E(u_t) = \frac{[\underline{n}] + [\bar{n}]}{2} \equiv \check{n}$$

and

$$\text{Var}(n_t) \equiv E(n_t - \check{n})^2 = \frac{1 - \rho}{1 + \rho} \text{Var}(u_t).$$

Thus, although our model has, strictly speaking, no stationary state (since u_t is uniformly distributed over $[[\underline{n}], [\bar{n}]]$), the dispersion remains constant around the mean value \check{n} . As a consequence, the general equilibrium model described below leads to a dynamic system which is ergodic with a bounded support. This will allow us to log-linearize the dynamic system around the average equilibrium defined by the mean value \check{n} , and will therefore make our model comparable to other existing business cycle models which have been solved and evaluated using similar methods.

Log-linearization results in a three-dimensional dynamic system of the following form (the hat on a variable denoting percentage point deviation from the average equilibrium):

$$\begin{bmatrix} \widehat{K}_t \\ \widehat{C}_t \\ \widehat{n}_t \end{bmatrix} = \Phi \begin{bmatrix} \widehat{K}_{t+1} \\ \widehat{C}_{t+1} \\ \widehat{n}_{t+1} \end{bmatrix} + \Gamma \begin{bmatrix} \widehat{e}_{t+1} \\ \widehat{u}_{t+1} \end{bmatrix}, \quad (25)$$

where $\widehat{e}_{t+1} = \widehat{C}_{t+1} - E_t(\widehat{C}_{t+1})$ is the one-step-ahead forecast error of consumption, and \widehat{u}_{t+1} is the ‘animal spirits’ innovation. Of course, all the other endogenous variables Y , H , w , r and μ can be uncovered by simple linear functions of the ‘state’ variables K , C and n .

The analysis of the matrix Φ is crucial for determining the properties of the average equilibrium. The three-dimensional dynamic system (25) includes two predetermined variables, \widehat{K} and \widehat{n} , and one non-predetermined variable, \widehat{C} . Hence, if the matrix Φ contains as many eigenvalues with modulus strictly above one as there are predetermined variables, the average equilibrium has a saddle-path property. As Benhabib and Farmer (1994) emphasized, this implies that the forecast error \widehat{e}_t is then a deterministic function of the endogenous disturbance \widehat{u}_t . If, on the contrary, all the eigenvalues of the matrix Φ have modulus strictly above one, the average equilibrium is a sink, and the forecast error may then enter as an independent shock to the business cycle. The properties of the matrix Φ depend of course

crucially on the specific values which are attributed to the structural parameters. In our simulation experiments, and for the reference calibration that we adopt below, we found that only two of the eigenvalues of the matrix Φ have modulus strictly above one, so that the average equilibrium around which we study the local dynamics has a saddle-path property.

4.1 Calibration

Calibration of most structural parameters is done within the range of admissible values which have been used in the related literature. Following Farmer and Guo (1994), we assume a discount factor $\beta = 0.99$, an average value of hours worked equal to 0.2 (20% of time endowment), a quarterly depreciation rate of capital $\delta = 0.025$, an elasticity of production with respect to capital $\alpha = 0.3$, and an infinitely elastic labor supply $\chi = 0$. Typical estimates of the inverse of the intertemporal elasticity of substitution range between 0 and 5. We set $\sigma = 2$, which is an average value not far from the logarithmic case. As for the parameter ρ , the threshold below which the intensity of the idiosyncratic signal is not sufficient to trigger a state transition in producers' conjectures, we assume $\rho = 0.9$ in order to get enough persistence.

Finally, a crucial parameter in our model is ϕ , the real fixed cost in the individual production function. This parameter is important for two reasons. First, it determines the admissible interval $[\underline{n}, \bar{n}]$ within which the number of active firms must remain, and second, it controls the degree of overall increasing returns (the inverse of the elasticity of the cost function) we are willing to introduce in our model:

$$\gamma = \frac{y + \phi}{y}.$$

There has been a great deal of empirical literature in recent years that has attempted to estimate the importance of returns to scale in actual economies. Since the earlier works by Hall (1990) and Caballero and Lyons (1992), who reported large IRS in the US manufacturing industry, recent studies by Basu and Fernald (1995, 1997) and Burnside, Eichenbaum and Rebelo (1995) have arrived at much smaller estimates, not very far from the assumption of constant returns to scale. However, even those recent papers include differences which, from the perspective of our paper, are particularly interesting. Indeed, while Basu and Fernald (1997) end up with nearly constant returns to scale for gross output as a whole, Burnside, Eichenbaum and Rebelo (1995)

find evidence of constant returns to scale with respect to *variable* inputs. As emphasized by Basu (1995) in his comment, the test of BER (1995) is actually a test to see whether firms operate under constant marginal costs. But if there are fixed costs in production activities, then BER’s estimates are fully consistent with globally increasing returns to scale, provided they appear under the mild form assumed in equation (13). Building on BER’s work and adding independent estimates of overhead costs provided by Ramey (1991), Basu (1995) shows that BER’s estimates are in fact consistent with 28% of increasing returns to scale.

In our simulations, we have chosen $\gamma = 1.23$ as a benchmark calibration, implying that fixed-cost represents 23% of individual net output, and resulting in the admissible interval $[3, 5]$ for the number of active firms. Clearly, given the recent available evidence, we see this calibration as ranging close to the upper bound of admissible values of increasing returns to scale. For this reason, we have also considered a second calibration with a much more conservative value for γ , and tested the robustness of our results against this alternative setting. Specifically, we assume in this second experiment that $\gamma = 1.05$, the value suggested by Basu (2002) to calibrate a production function similar to ours. For this value, a simple computation shows that the admissible interval for the number of active firms is $[11, 21]$.

4.2 Empirical properties

We now compute the second order moments for the main variables of our model, and compare them to their empirical counterparts.¹³ Tables 1 and 2 report the relative volatilities and cross-correlations with output of the main variables as measured in US data, and as implied by our model, under the two different assumptions regarding the size of increasing returns to scale.

The most obvious observation from these tables is that the kind of business cycles implied by the model resemble closely that of the US economy. In particular, Table 1 shows that for the two calibrations considered, the model easily reproduces the weaker volatilities of consumption and real wages relative to output, and the larger volatility of investment. The volatility of hours

¹³To compute the theoretical moments, we generated 100 series of 150 innovations uniformly taken on the support $\{[\underline{n}], [\bar{n}]\}$ and constructed the series for all endogenous variables. The average moments are reported in Table 1, after the series have been filtered by the Hodrick-Prescott filter. Empirical moments are taken from King and Rebelo (1999).

Table 1 - Absolute and relative standard deviations

$\frac{\sigma_Y}{Y}$	Variable (X)	$\frac{\sigma_X}{\sigma_Y}$			
		C	H	I	w
1.81	US Data	0.74	0.99	2.93	0.38
1.79	Model ($\gamma = 1.23$)	0.62	1.54	2.57	0.08
0.35	Model ($\gamma = 1.05$)	0.71	1.76	2.13	0.09

is somewhat larger than in the data, but this is a natural feature of models which do not introduce external effects on labor or exogenous shocks to factor productivity. In addition, Table 2 shows that the model also accounts for the strong procyclical movements of consumption, employment and investment that are typical of real economies.

To understand how fluctuations in the equilibrium number of active firms generate such fluctuations in real variables, it is useful to consider the pseudo impulse-response functions to a positive “animal spirits” shock that we displayed in Figure 3. Figure 3 plots the expected behavior of the main variables (expressed as percentage point deviations from the average state) when it is assumed that the current realization of u_t is 5 and all future values are 4, the mean value of the admissible interval $[[\underline{n}], [\bar{n}]]$. Note that given the calibrated value of $\rho = 0.9$, this realization of u_t implies that the average number of firms increases by only 2.5% (i.e., raises from 4 to 4.1). Following the entry of new firms, there is an immediate and persistent reduction in the average markup, which generates in turn an instantaneous and sustained increase in the demand addressed to each sector. Firms meet this increased demand by raising output, investment and employment. With a labor demand curve shifting along an almost invariant horizontal labor supply curve (remember that the wage-elasticity of hours is infinite), real wages also increase during the boom, albeit only slightly during the first periods. With consumers working more at a higher wage, consumption also rises persistently.

Obtaining simultaneous procyclical movements of output, consumption and investment is particularly important in our model since, as pointed out by Benhabib and Farmer (1999), standard models of indeterminacy via a sink stationary equilibrium have encountered significant difficulties in explaining this feature without relying on large markups or large increasing returns

Table 2 - Contemporaneous correlations with output

Variable (X)	$Corr(Y_t, X_t)$			
	C	H	I	w
US data	0.88	0.88	0.90	0.12
Model ($\gamma = 1.23$)	0.99	0.99	0.99	0.19
Model ($\gamma = 1.05$)	0.99	0.96	0.99	-0.60

through declining marginal costs.¹⁴ This is because with a constant demand for labor and conventional slopes for the labor demand and labor supply curves, any shift in the labor supply schedule necessarily implies countercyclical movements of the real wage, and thus generally leads to decreased consumption.

In that respect, a key feature of our model is that positive “animal spirits” (leading to the entry of new active firms) induce a decrease in the average markup and thus an outward shift in the demand for labor. This increase in labor demand may be strong enough to offset the depressing effect of labor productivity on real wages, so that real wages eventually rise in response to a positive sunspot. Although this simple mechanism is similar to the countercyclical markup models of Galí (1994) and Schmitt-Grohé (1997), the main obvious difference is that we do not have to require extremely large markups or increasing returns to make them fluctuate endogenously.¹⁵ In fact, as shown in Table 2, the model still generates procyclical movements of consumption with 5% of increasing returns ($\gamma = 1.05$).

Hence, by the standards of the business cycle literature, the model does pretty well at accounting for the main features of US postwar fluctuations. Interestingly, most of these results are robust to the different calibrations on the degree of returns to scale γ . In fact, as shown in Table 1, the only signifi-

¹⁴See amongst others Benhabib and Farmer (1994), Farmer and Guo (1994), Benhabib and Farmer (1996), and Schmitt-Grohé (2000) for additional details. An interesting exception is the two sector model with home production of Perli (1998), in which an “animal spirits” shock both increases output and market consumption, because of an (unobserved) fall in home production. However, calibrated versions of this model imply that the degree of increasing returns to scale must be set to 20% in order to generate procyclical consumption and sufficient autocorrelation in output.

¹⁵For example, in Galí (1994), values for the markup above 2.0 are necessary to obtain an indeterminate equilibrium. Galí even uses a markup of 2.8 to generate good results in his simulations.

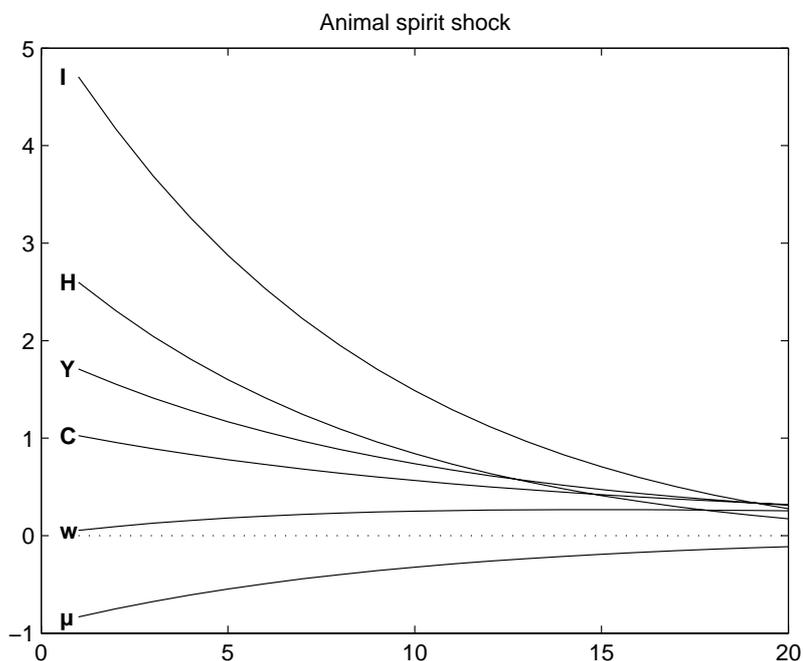


Figure 3: Pseudo impulse response functions for the benchmark model.

cant differences concern the contemporaneous correlation of output with the real wage (with $\gamma = 1.05$, the real wage decreases in the first period before rising up, so that the instantaneous correlation is negative) and the absolute volatility of output. With $\gamma = 1.23$, this volatility is similar to that observed in the US economy, which might be considered as implausibly high for a model with a unique source of disturbances. With $\gamma = 1.05$, this variance is still around 20% that of the US economy. Hence, considering large or small degrees of increasing returns to scale mainly influences the potential of the mechanism we emphasized in this paper to account for a significant part of actual output fluctuations.

5 Conclusion

The idea that endogenous changes in the beliefs of economic players may influence the level of economic activity has been revived by the recent research on business cycles. In particular, dynamic general equilibrium models with a sink stationary equilibrium have shown that endogenous changes in beliefs can be self-fulfilling, and hence consistent with individual optimization and rational expectations.

In this paper, we have explored another way of generating such endogenously driven fluctuations by exploiting the idea, put forward in the microeconomic literature, that different levels of economic activity sustained by different numbers of active firms may be consistent with conditions of free entry and perfect contestability within each sector. We have shown that a simple economy displaying this kind of indeterminacy may be subject to large fluctuations due to endogenous changes in the animal spirits of firms and in their willingness to enter and compete within existing markets. Furthermore, these fluctuations have properties similar to those observed in the US economy.

Finally, because our model does not rely on the sink property of the steady state to generate indeterminacy, it avoids some difficulties that are encountered in recent related models, such as the requirement either to impose large increasing returns through declining marginal costs, or to trade lower increasing returns against countercyclical movements of consumption.

References

- d'Aspremont, C., Dos Santos Ferreira, R. and L.-A. Gérard-Varet, 2000, Contestability and the indeterminacy of free entry equilibria, *Japanese Economic Review* 51, 72-83.
- Basu, S., 1995, Capacity utilization and returns to scale - Comment, in: S. Bernanke and J.J. Rotemberg, eds., *NBER Macroeconomics Annual* 10, 110-118.
- Basu, S., 2002, Technology and business cycles: How well do standard models explain the facts?, in: J. Fuhrer and S. Schuh, eds., *Beyonds shocks: What cause business cycles?*, Conference Proceedings, Federal Reserve Bank of Boston.
- Basu, S. and J. G. Fernald, 1995, Are apparent productive spillovers a figment of specification error?, *Journal of Monetary Economics* 36, 165-188.
- Basu, S., and J. G. Fernald, 1997, Returns to scale in U.S. production: Estimates and implications, *Journal of Political Economy* 105, 249-283.
- Benhabib, J. and R. Farmer, 1994, Indeterminacy and increasing returns, *Journal of Economic Theory* 63, 19-41.
- Benhabib, J. and R. Farmer, 1996, Indeterminacy and sector specific externalities, *Journal of Monetary Economics* 37, 397-419.
- Benhabib, J., and R. Farmer, 1999, Indeterminacy and sunspots in macroeconomics, in: J.B. Taylor and M. Woodford, eds., *Handbook of macroeconomics*, Vol. 1 (North-Holland, Amsterdam) 387-448.
- Berchtold, A., 1999, The Double Chain Markov model, *Communications in Statistics: Theory and Methods* 28, 2569-2589.
- Brock, W. A. and J. A. Scheinkman, 1983, Free entry and the sustainability of natural monopoly: Bertrand revisited by Cournot, in: D.S. Evans, ed., *Essays on Industrial Organization and Regulation* (North-Holland, Amsterdam) 231-252.

- Burnside, C., Eichenbaum, M. and S. T. Rebelo, 1995, Capacity utilization and returns to scale, in: S. Bernanke and J.J. Rotemberg, eds., NBER Macroeconomics Annual 10, 67-110.
- Caballero, R. J., 1992, A fallacy of composition, *American Economic Review* 82, 1279-1292.
- Caballero, R. J. and R. K. Lyons, 1992, External effects in U.S. procyclical productivity, *Journal of Monetary Economics* 29, 209-225.
- Chatterjee, S. and R. Cooper, 1993, Entry and exit, product variety and the business cycle, NBER Working Paper #4562.
- Chatterjee, S., Cooper, R. and B. Ravikumar, 1993, Strategic complementarity in business formation: Aggregate fluctuations and sunspot equilibria, *Review of Economic Studies* 60, 795-811.
- Farmer, R. and J. T. Guo, 1994, Real business cycles and the animal spirits hypothesis, *Journal of Economic Theory* 63, 42-72.
- Galí, J., 1994, Monopolistic competition, business cycles, and the composition of aggregate demand, *Journal of Economic Theory* 63, 73-96.
- Hall, R., 1990, Invariance properties of Solow's productivity residual, in: P. Diamond, ed., *Growth, Productivity, Employment* (MIT Press, Cambridge), 71-112.
- King, R. G. and S. Rebelo, 1999, Resuscitating Real Business Cycles, in: J.B. Taylor and M. Woodford, eds., *Handbook of macroeconomics*, Vol. 1 (North-Holland, Amsterdam), 927-1007.
- Hansen, G., 1985, Indivisible labor and the business cycle, *Journal of Monetary Economics* 16, 309-327.
- Knieps, G. and I. Vogelsang, 1982, The sustainability concept under alternative behavioral assumptions, *Bell Journal of Economics* 13, 234-241.
- Novshek, W., 1980, Cournot equilibrium with free entry, *Review of Economic Studies* 47, 473-486.
- Perli, R., 1998, Indeterminacy, home production and the business cycle: A calibrated analysis, *Journal of Monetary Economics* 41, 105-125.

- Portier, F., 1995, Business formation and cyclical markups in the French business cycle, *Annales d'Economie et de Statistique* 37/38, 411-440.
- Ramey, V. A., 1991, Comment, in: O.J. Blanchard and S. Fisher, eds., *NBER Macroeconomics Annual* 6, 141-145.
- Rotemberg, J. J. and M. Woodford, 1995, Dynamic general equilibrium models with imperfectly competitive product markets, in: T.F. Cooley, ed., *Frontiers of Business Cycles Research* (Princeton University Press), 243-293.
- Schmitt-Grohé, S., 1997, Comparing four models of aggregate fluctuations due to self-fulfilling expectations, *Journal of Economic Theory* 72, 96-147.
- Schmitt-Grohé, S., 2000, Endogenous business cycles and the dynamics of output, hours and consumption, *American Economic Review* 90, 1136-1159.
- Spitzer, F., 1976, *Principles of Random Walk* (Springer-Verlag).
- Weder, M., 2000, Animal spirits, technology shocks and the business cycle, *Journal of Economic Dynamics and Control* 24, 273-295.
- Wen, Y., 1998, Capacity utilization under increasing returns to scale, *Journal of Economic Theory* 81, 7-36.