Machine Replacement, Network Externalities, and Investment Cycles*

Juan M. Ruiz†
Universidad Carlos III de Madrid

July 4, 2000

Abstract

This paper presents a model where agents decide on the timing of replacement of ageing machines. The optimal replacement policy for an agent is influenced by other agents' decisions because the productivity of a particular vintage displays network externalities that set in with a lag. In equilibrium, agents follow innovation cycles with a frequency that is lower than optimal, so there is too much delay. One extreme case is the possibility of inefficient collapse: for some parameters there is no investment in equilibrium, even though it is socially optimal that agents (eventually) invest in cycles. Another feature of the model is the tendency of agents to synchronize their individual decisions, and thus the outcome of the aggregate economy does not smooth out the non-convexities present at the microeconomic level.

(J.E.L. Classification Numbers D92, E22, E32)

*Updated versions can be downloaded at `http://www.eco.uc3m.es/~jruiz/machdis.pdf`

This paper grew out of discussions with my advisors Russell Cooper and Douglas Gale, to whom I am immensely grateful for their comments and guidance. I have also benefited from the insighful comments of Samuel Kortum, Debraj Ray, Matilde Machado and seminar participants at the XV Latin American Meeting of the Econometric Society, NBER Summer Institute and Universities of Alicante, Carlos III (Madrid), Navarra, Nova (Lisbon), Pompeu Fabra, and Boston University. Financial support from Boston University is gratefully aknowledged. Needless to say, all remaining errors are mine.

†**Mailing Address:** Universidad Carlos III de Madrid, Departament of Economics, 28903 Getafe (Madrid), SPAIN. Tel.: +(34) 91 624 9652. Fax : +(34) 91 624 9875. e-mail: `jruiz@bu.edu`

1 Introduction

This paper explores the dynamics of macroeconomic aggregates when underlying microeconomic decisions are non-convex and network externalities affect the profitability of each agents' action with a lag. We ask two specific questions. First, do agents coordinate their actions in the presence of network externalities, and thus transmit the variability of their micro actions into aggregate variables; and second, is the social optimum the outcome of one of the equilibria in this setting. We find that, in equilibrium, agents tend to synchronize their actions, and thus if we look at an aggregate variable, we will still observe some lumpiness that is not smoothed out by aggregation. Moreover, these aggregate cycles display a frequency that is lower than optimal, so for a number of users bigger than a specific threshold, there is *always* inefficient delay in this economy, whereas in previous work the outcome of one of the (multiple) equilibria was equivalent to the socially optimal policy.

The issue of whether the lumpiness of economic decisions at the microeconomic level gets smoothed out by aggregation is not a theoretical curiosity. Work by Bertola and Caballero (1990) and Cooper, Haltiwanger and Power (1995) for example, show that the presence of nonconvexities at the microeconomic level is an important feature that we have to consider if we want to explain the lumpiness found in the data, especially for investment in equipment and purchases of durable goods. It is also this lumpiness at the micro level (usually generated by some degree of irreversibility), together with the evolution of the cross sectional distribution of agents, that may generate interesting aggregate dynamics when coupled with idiosyncratic and common uncertainty.

The issue of externalities and the timing of individual decisions has been extensively studied in the literature. Previous work by Shleifer (1986) and Gale (1995, 1996) showed how the presence of positive externalities can lead to delay of individual's actions and to cycles in macroeconomic aggregates. In Shleifer (1986), externalities are pecuniary in nature, and the continuous arrival of investment opportunities is transformed into aggregate cycles of investment activity. This is the result of the incentive to synchronize individual investment decisions to take advantage of the positive externality through increased aggregate activity. While Shleifer emphasizes the multiplicity of equilibria in this context, our work focuses on the amplification of the nonconvexities observed at the individual level and on the efficiency properties of the equilibrium cycles.

Gale (1995) focuses on the incentives for delay in a dynamic model of coordination. The basic idea is that agents would want to "invest" at times of high aggregate activity, and this coordination problem introduces the possibility of inefficient delay. One particular shortcoming of this model is that, even though the game has an infinite horizon, agents get to act only once, so some interesting issues arising in investment decisions (which usually call for agents to act more than once) cannot be analyzed. Gale (1996) develops a model addressing precisely that issue, with agents that get to act more that once in the course of the game.[1] In this second paper, there is again a positive externality from the level of aggregate activity, which generates incentives for delay in recessions, effectively increasing the amplitude and reducing the frequency of the underlying (stochastic) cycle.

Another strand of the literature has studied the effect of complementarities on the behavior of macroeconomic aggregates. The closest work to our own is that of Cooper and Haltiwanger (1992), where it is shown that macroeconomic aggregates do not necessarily smooth out non-convexities at the individual level. The basic idea is that the presence of strategic complementarities gives an incentive to individual agents to "bunch" their actions, and so the lumpiness intrinsic in the individual decision process is transmitted into the aggregate time series of investment. The analysis is carried out separately for the timing decision (synchronization versus staggering), and then for the periodicity of the action. However, when both the timing and periodicity decisions are taken together, the problem of multiplicity of equilibria makes it very difficult to give a clear prediction about the outcome one would expect from the model.

This paper attempts to combine the two strands of the literature on externalities and the timing of individual decisions, and analyze the effect of complementarities on the behavior of macroeconomic aggregates, when individuals face a non-convex problem. In particular, we consider a model where agents are infinitely lived and each one operates a machine that is subject to obsolescence due to the arrival of new vintages every period. At any time, each agent has the option of switching to another vintage after incurring a fixed cost. This describes a well known machine replacement problem that has been extensively studied in the operations research literature. In this deterministic setup, the optimal policy calls for an agent to update his machine to the "state of the art" machine at regular intervals. A complete characterization of the solution in this case is shown in the first section of the paper and it serves as a benchmark for the rest of the analysis.

In order to introduce network externalities into the problem, we assume that the profitability of using a particular vintage depends on the number of users of that vintage. As it is well known, and the papers mentioned above suggest, one of the problems of models with payoff spillovers is the emergence of multiple equilibria, especially when agents are forward looking.[2] However, some

[1] Although the results in Gale (1996) require that an each period there is a positive probability that an agent will exit the game forever, so in effect agents in this model are not infinitely lived.

[2] See for example Krugman (1991) and Matsuyama (1991) for a discussion of the role of expectations in models

recent literature has focused on the *timing* of the externality (i.e. contemporaneous versus leading or lagging externalities) and the effect on the multiplicity of equilibria.[3] Following that literature, we introduce externalities that set in with a lag, so that the profitability of using a particular vintage depends on the number of people who used that vintage the previous period. For new machines that just appeared in the market, there is no history of previous use, and so, no externality built-in. Thus, two forces influence the returns from a particular vintage in the first periods of its appearance in the market: obsolescence reduces the profitability of a machine as time passes, whereas, if agents coordinate on adopting a particular vintage, when the externality sets in, it may well outweigh the effect of obsolescence at least initially.

The introduction of lagging externalities gives rise to interesting phenomena in this model that distinguish it from previous work, especially Shleifer (1986) and Cooper and Haltiwanger (1992). First of all, it greatly reduces the set of equilibria (as compared to the case of contemporaneous externalities), although, contrary to Adserà and Ray (1997), it does not result in a unique equilibrium except for very particular cases. Second, contrary to Shleifer (1986) and Cooper and Haltiwanger (1992), when the user base is big enough, all equilibria in our model are inefficient, since the innovation cycles that arise in equilibrium have a lower frequency than optimal. Our model exhibits inefficient delay since there is a tendency for agents to free ride on other agents adopting a new technology first, (delaying the switching cost of adopting a new machine and leaving to others the task of building up the externality associated with the new adopted vintage), only to follow them after a few periods and switch to that vintage (even though it is no longer the "state of the art"), when it has been used for some time and the effect of network externalities has had time to "set in".

The last interesting result is derived from the incentive to coordinate the adoption of new machines to take advantage of the vintage-specific spillovers. This in turn implies that the aggregate time series of investment in this model will not smooth out the non-convexities found at the microeconomic level. Although we can construct a staggered equilibrium that would in practice provide the smoothing of individual lumpiness for any size of the user base, it does not seem to be robust to a small perturbation of the distribution of agents across vintages, and the synchronized equilibrium is shown to be the unique (stable) steady state for the particular case of zero switching costs.[4]

with forward looking, infinitely lived agents whose decisions have some degree of irreversibility and where there are payoff externalities.

[3] In Gale (1995), the issue of the timing of the externality is analyzed in discrete time and in a strategic context. In Gale (1996), lagging externalities help to pin down a unique equilibrium for the dynamic game. Using a setup very similar to that of Krugman (1991), Adserá and Ray (1997) show how imposing a lag for the externality to "set in" reduces the set of equilibria, and, with non-strategic agents, results in a *unique* equilibrium

[4] In a separate paper, we analyze the robustness of this result to the introduction of a small degree of heterogeneity on the distribution of agents operating machines. For example, when the adoption costs are different across consumers,

The next section describes the model in detail, specifying the particular form of network externalities introduced and the choice of solution concept. Section three develops the solution to the planner's problem, for a particular initial condition. Those results would then serve as a benchmark to analyze the efficiency properties of the equilibrium cycles. Section four deals with the conditions necessary for synchronization in the steady state and the efficiency properties of the equilibrium cycles. In section five we introduce staggered equilibria and give some intuition why the staggered steady state may be considered a "knife edge" case in a sense made specific in that section. Parts six and seven give further insight into why synchronization may be considered the "natural" outcome of the game. Finally, section eight include some concluding remarks and further avenues of research.

2 Model

This is an economy with a mass N of infinitesimal, infinitely-lived agents. Each agent operates one machine, indexed by age x (which is also the time elapsed since it appeared in the market). New machines appear every period and agents have at any time the possibility to scrap their machine and adopt a machine of a different vintage (age) by paying the switching cost $c \geq 0$. The problem of each agent is to choose switching dates and target vintages over an infinite horizon to maximize the present discounted value of payoffs net of switching costs. Each agent discounts the future at the common discount factor $\beta = \frac{1}{1+r}$.

2.1 Payoffs

At each period t, the pattern of use of different vintages across agents can be characterized by the distribution $n_t = \{n_{0t}, n_{1t}, n_{2t}, \ldots\}$, where n_{xt} represents the mass of users of vintage x at time t and of course we require $0 \leq n_{xt} \leq N$ and $\sum_{x=0}^{\infty} n_{xt} = N$ at every t.

To introduce lagging externalities into the model, we will assume that the payoff derived from a machine of a particular vintage depends, in part, on how extensively it was used the *previous* period, and will denote by $F(x, n_{x-1,t-1})$ the payoff at time t of using a machine of age x. The motivation for a payoff function of this sort could be any industry which depends on another "upstream" industry for "technical support" for example. We may think that the availability of support —the number of repair shops, for example— for a particular vintage depends on how spread the use of that vintage

the synchronized equilibrium will still be an equilibrium in the modified model when the variability of those costs is small enough.

is, which would provide the network effect. With respect to the timing of the externality, it is very likely that this "support industry" would exhibit some inertia in adjusting to changes in the industry it supports: it would need some time to develop after the users of a particular vintage decide to switch to a particular brand, and also it will need some time for that support to disappear even after a significant reduction in the number of users, so we would have a two-sided lag.[5]

It is worth mentioning that the qualitative results of the paper —the presence of inefficient delay and the transmission of the micro non-convexities into the aggregate data— do not change with the length of the time period, so in this context, the externality lag could be very small indeed.[6] Moreover, taking the approach of specifying a one-period-lag for the appearance and disappearance of the externalities is just a simple case of a general distributed lag scheme including possibly the current user base.[7]

Ignoring integer constraints on the age of the machine x, we will assume that[8]

$$F(x,n) > 0 \; \forall x, n \tag{1}$$

$$F_1 < 0 \; \forall x \geq 1 \tag{2}$$

$$F_{11} > 0 \; \forall x \geq 1 \tag{3}$$

$$F_2 > 0 \; \forall x \geq 1 \tag{4}$$

[5] If we take investment in durable goods for example, a real life example could be the availability of repair shops specialized in a particular car brand, or even the availability of spare parts for a particular model. Presumably those shops appear in response to the number of users of a particular brand or model, and it takes some time for them to react to the change in the user base of those models. These inertia could be also the result of uncertainty on government policy: in the case of Colombia or Peru, for example, repair shops and spare parts distributors sometimes are reluctant to enter into the market at the same time a new car brand starts to be imported into the country, because of fear of changes in import restrictions, for example.

[6] In Ruiz (1998) we also analyze the effects of considering a "stock" version of this externalities: we could think of payoff externalities because of learning by using, and so the aggregate number of users in the past determines the profitability of a particular vintage. In that case, even when everyone abandons a model, the payoff of that machine does not drop, since the stock of knowledge associated with that machine is the same. The qualitative results are essentially the same in that setting: inefficient delay and non-smoothing by aggregation.

[7] The qualitative results of the paper do not change by introducing a general distributed lag of user base in the determination of the network externalities. Including the current user base in the distributed lag could only affect the result of the inefficiency of the equilibrium cycles. However, if the weight put on the current consumer base is low enough, as compared with the past user base, the results of this paper are still valid.

[8] Strictly speaking, we should replace (2) by

$$F(x_1, n) > F(x_2, n) \text{ if } 1 \leq x_1 < x_2,$$

(3) by

$$F(x_1, n) - F(x_1 + k, n) > F(x_2, n) - F(x_2 + k, n) \text{ if } 1 \leq x_1 < x_2,$$

and (5) by

$$F(x_1, n_1) - F(x_1 + k, n_1) > F(x_1, n_2) - F(x_1 + k, n_2) \text{ if } x \geq 1 \text{ and } n_1 < n_2$$

$$F_{21} \quad < \quad 0 \; \forall x \geq 1 \tag{5}$$

Assumption (1) implies that it is always profitable to use a machine, regardless of the number of users of the same vintage. This in fact allows us to avoid the issue of exit and re-entry which could greatly complicate the model without any gain in insight. Assumption (2) represents the effect of obsolescence on the profitability of a particular vintage, which can be understood as a normalization with respect to the "state of the art" technology,[9] whereas the effect of positive network externalities is captured by (4). We think of the effect of obsolescence as reducing current profitability by a fixed fraction each period a new machine arrives. That motivates assumption (3), and together with network externalities, assumption (5). Another way to interpret condition (5) is to think of weaker network externalities for older machines.

We will also assume that if a machine is more profitable than another, then the effect of obsolescence does not modify this relative ranking: the former would still be more profitable in the future, provided there is no change in the number of users of both of them:

$$F(x_1, n_1) > F(x_2, n_2) \Rightarrow F(x_1 + k, n_1) > F(x_2 + k, n_2) \tag{6}$$

Notice that the payoff of a new machine ($x = 0$) is still not well defined by these assumptions on $F(.)$. Since the network externalities we are dealing with appear with a lag, the externalities associated with a *new* machine have not had time to be built into the vintage. So the number of agents using a new machine is irrelevant for the *current* profitability of that vintage (although it will influence its profitability next period). Accordingly, we define

$$F(0, n) \equiv F(0, 0) = 1 \text{ for all } n \tag{7}$$

where the last equality is just a convenient normalization of units and is without loss of generality.

It is worth noting that, for given n, the function $F(x, n)$ is the result of two opposing forces: on the one hand, there is the negative effect of obsolescence as a machine gets older, captured by $F_1 < 0$. On the other hand, an old machine may be more productive than the "state of the art"

[9] In principle we can think of an alternative model in which the profitability of a machine is only affected by physical depreciation but not by the appearance of a more productive vintage. However, we can always turn that model into an equivalent model where the payoff of each machine is normalized with respect to the profitability of the "state of the art" technology appearing each period as in our case.

The discount factor β in the normalized model would incorporate not only the effect of discounting in the original model, but also the effect of the speed of technological progress, which we would require to be not too high with respect to the original discount factor. For an example of this kind of normalization, see Cooper, Haltiwanger and Power (1995). The advantage of using this normalization is that the model becomes stationary, with the consequent gain in tractability.

technology, if it has been used in the past by a large enough mass of agents. It is possible, then, that the total number of agents N who own a machine satisfies

$$F(1, 0) < 1 = F(0, n) < F(1, N) \tag{8}$$

In general, we call n_l the number of users that makes a machine as profitable in the second period of use (age = 1) as a new one, that is

$$F(1, n_l) = 1 \tag{9}$$

In other words, for a vintage used *continuously* since its appearance in the market by a mass of users bigger than n_l, the payoff is increasing from the first period to the second and it is decreasing thereafter (see figure 1). In the sequel, this need to "build up" the externality associated with a particular vintage will be the source of inefficient delay, as agents will want to wait for others to bear the cost of building the user base, only to join them later.

Notice that, given assumption (7) the function $F(x, n)$ is not monotonic in x for $N > n_l$ (it would be monotonic only if $N \leq n_l$), and so an old machine with a user base bigger than n_l may still provide a higher per period payoff than a new machine. In the sequel, it will be important to refer to the maximum age of a machine for which this is still true. We thus define $\hat{t}_N$ as that limit age:

$$\hat{t}_N = Max \{t : F(t, N) \geq F(0, N) = 1\} \tag{10}$$

So, for example, t_N is equal to zero whenever $N < n_l$, as in figure 1, it is equal to one for $N = n_l$ and it is greater or equal to one for $N > n_l$. In the case of the figure, $t_N = 2$ for the payoff function with the highest mass of users (the topmost curve).

An example of a particular functional form that satisfies (1) through (6) is $F(x, n) = \delta^x(n + 1)$ for $x \geq 1$, where δ is the fixed obsolescence factor. For this functional form, $n_l = \frac{1-\delta}{\delta}$ and (ignoring integer constraints) $\hat{t}_n = -\frac{\log(n+1)}{\log \delta}$.

2.2 Strategies and Equilibrium Concept

Each agent is endowed with a machine at time $t = 0$. A strategy for agent i is a mapping $s_i : \mathcal{N} \to \mathcal{N}$, where $s_i(t)$ indicates the age of the machine used by agent i at time t and $\mathcal{N}$ represents the set of natural numbers. Notice that we use open-loop strategies, that is, strategies as a function of time only, and not as a function of the whole history of the game up to time t. The equilibrium concept we will use is **open-loop equilibrium**, that is, Nash equilibrium in open loop strategies.

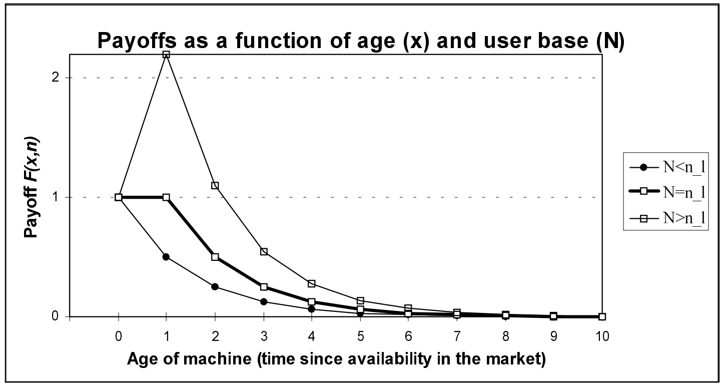

Figure 1: Per-period payoff as a function of the age x of the machine for a given (and fixed) number of users N. (n_l stands for n_l in the text)

The use of open-loop strategies instead of closed-loop strategies implies that agents may not be using all relevant information to define their optimal strategies. In particular, they do not react to deviations of a single agent. However, this is not a restrictive assumption in this setting since we have assumed infinitesimal players and payoffs that are independent of deviations of a subset of agents of zero measure, so agents can "safely" ignore those deviations. With these two conditions (infinitesimal players, and payoffs independent of single-agent deviations), the *outcome* of an open-loop equilibrium is subgame perfect, although open-loop *strategies* may not be perfect[10]. The use of open-loop strategies has also the advantage of reducing considerably the strategy space, with the added gain in tractability.

3 Planner's Solution

As a useful benchmark to which we can compare the equilibrium outcome of the game described above, we examine the centralized solution in this setting. Suppose there is a planner who starts with a mass N of machines of the *same* vintage, or using the notation before, $n_{i0} = N$ for some i. The objective of the planner is to maximize the sum of individual payoffs (net of switching costs) derived from these machines, by choosing the adoption dates (and vintages to be adopted at such

[10] For a discussion of this point, see for example Fudenberg and Tirole (1991).

adoption dates) for each machine. In order to tackle this multi-machine problem, we will follow two steps. First, we consider the solution of a planner who is forced to choose always the same vintage for all his machines. Second, we examine whether this restricted solution would change by allowing the planner to split his machines into two or more vintages at any point in time.

3.1 Restricted planner: only one vintage each period

In this first case we solve the problem of a planner who cannot hold machines of more than one vintage at a time. Since operating a machine is always profitable, this planner will always operate N machines of the same vintage, and we can reduce the problem to that of maximizing the flow of payoffs of a representative machine. Taking c as the cost of switching each machine, the planner's problem becomes

$$\underset{\{s_t\}}{Max}\ N\left\{\sum_{t=0}^{\infty}\beta^t\left[F\left(s_t,N\right)-cz_t\right]\right\} \quad (11)$$

$$z_t = \begin{cases} 0 & \text{if } s_t = s_{t-1}+1 \\ 1 & \text{otherwise} \end{cases} \quad (12)$$

where s_t is the vintage chosen at period t and z_t is a binary variable that takes the value of one when there is a switch in technology.

In principle, the planner has the option of switching to *any* vintage at every period t. However, since the switching cost c is the same regardless of the vintage it chooses to switch to, the only relevant decision at any moment is whether to keep his machines or to update them to the state of the art technology. This means that we can restrict attention to the choice $s_t \in \{0, s_{t-1}+1\}$.[11]

Taking advantage of the stationary nature of the problem, we can rewrite it in a recursive way. The present discounted value (per machine) of having a group of machines of age x synchronized with a mass N of machines) is then given by the value function $V(x;N)$:

$$V(x;N) = Max\left\{-c+F(0,N)+\beta V(1;N), F(x,N)+\beta V(x+1;N)\right\} \quad (13)$$

where the first term in the maximization problem represents the gains from switching to a new machine of age $x=0$: the planner pays the fixed cost c per machine and gets the returns from having a machine of age 0 this period and next period it starts with a machine of age 1, which has a value $V(1)$. The second term represents the gains from choosing not to innovate, receiving the payoff of a machine of age x and starting the next period with a machine of age $x+1$. It is

[11] However, because of the presence of network externalities, this is not true for individual agents, as will become evident when we analyze the decentralized outcome to the model.

easy to show that $V(x;N)$ is increasing in the user base N, because of assumption (4) about the instantaneous payoff function $F(.)$. Moreover, the value function $V(x;N)$ is decreasing in the age x of the machines over the range $x \geq 1$.[12]

The optimal replacement rule for this problem has been extensively studied in the operations research literature. As one would expect, the optimal policy is a stationary control-limit rule specifying a threshold $i_p(N)$ such that the machines should be replaced if and only if their age x is higher than this threshold, that is $x_t = \varphi(x_{t-1})$, where the policy function $\varphi(.)$ satisfies:

$$\varphi(x') = \begin{cases} x'+1 & \text{if } x' \leq i_p(N) \\ 0 & \text{if } x' > i_p(N) \end{cases} \tag{14}$$

where $i_p(N)$ is the minimum integer i solving (Cf. Beckmann (1968)):

$$c \leq \sum_{x=0}^{i} \beta^x \left[F(x,N) - F(i,N)\right] \tag{15}$$

and the second order condition requires $F_1(i_p(N),N) < 0$ which is satisfied by the assumptions on $F(.)$. Manipulating (15) gives

$$c + \frac{F(i_p+1,n)}{1-\beta} \leq \frac{-c\beta^{i_p+1} + \sum_{k=0}^{i_p} \beta^k F(k,N)}{1-\beta^{i_p+1}} \equiv V(0;N,i_p+1) \tag{16}$$

which expresses the more intuitive condition that in order to find the optimal innovation threshold i_p, the marginal gain of using a longer cycle of length i_p+1 (the left side of the inequality) should be no higher than the average gain from that extra period in the adoption cycle, expressed in the right side of the inequality.[13]

This optimal policy effectively implies adoption cycles of constant length $i_p(n)$. It is clear from (15) that the length of these cycles is increasing in the switching cost c. However, the length of the optimal replacement cycle is not necessarily monotonic in N even when $F_{12} < 0$.[14] The reason is that, due to lagging externalities and assumption (7), the first difference $F(0,N) - F(i,N)$ is decreasing in N, whereas assumption (5) implies that the differences $F(x,N) - F(i,N)$ will increase with N for all the other terms of the summation.

[12]The usual monotonicity argument applies: a planner with a machine of age $x \geq 1$ can always replicate the policy of any other planner with machines of age $y \geq x$, and get a payoff that is no lower than that obtained by the latter, since $F(x;.)$ is decreasing for $x \geq 1$. Of course that is not necessarily true for $x = 0$ since $F(0,N) < F(x,N)$ for x low enough and N high enough, by assumption (8).

[13]Of course, the inequalities in (15) and (16) are due to time being discrete in this model. In an alternative model with continuous time the two conditions would hold with equality.

[14]In fact, it is easy to generate examples of F functions for which the optimal threshold $i_p(N)$ is increasing for low values of N and decreasing for high values of N.

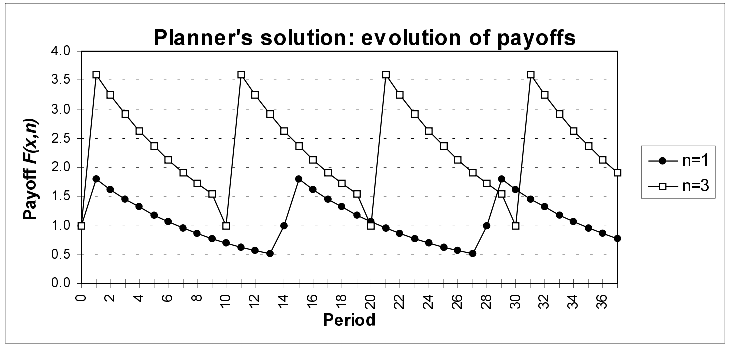

Figure 2: Evolution of planner's per period payoff for two different sizes of the mass n of machines.

3.2 Allowing simultaneous vintages

The previous exercise is the solution to a constrained problem since this planner is not allowed to split his machines into different vintages, starting from a bunched distribution. As opposed to the usual replacement problem studied in the operations research literature, in this case, there is also an externality problem. Although there is a positive network externality that would give incentives to the planner to keep all machines in the same vintage, it is also true that these machines have to bear the initial cost of "building" the externality associated with that vintage in the first period.

The trade off between delaying the cost of adoption for a fraction of the machines and the (positive) externality loss from "splitting" a homogeneous group of machines may result in the planner splitting his machines into more than one vintage for some periods, resulting in a higher *average* net profitability. To see why this may be so, consider a planner using threshold $i_p(N)$ as before. Suppose he chooses to replace only a fraction γ of his machines when they reach age $i_p(N)$ and the rest of them one period later, so that two periods ahead he again has a homogeneous group of machines of age 1. Compared to the previous solution, he saves $(1-\gamma)(1-\beta)c$ on the delayed switching cost for a fraction $(1-\gamma)$ of his machines. On the other hand, he suffers a reduction in profitability during two periods: in the first period, the loss equals $(1-\gamma)\left[1 - F(i_p+1,N)\right]$ from the group which fails to innovate to a (possibly) more profitable machine,[15] and in the second

[15]Notice that the solution $i_p(N)$ to the constrained planner problem may involve innovating to a new machine that has an initial *lower* productivity than the currently owned, that is, it may be that $F(i_p,N) > F(0,N) \equiv 1$. In this case (which requires c large, this effect would give more incentives to the planner to split his machines into more than one group.

period the loss is equal to $F(1,N) - F(1-\gamma N)$ which reflects that the machine was adopted only by a fraction γ of users when it appeared in the market, thereby reducing its profitability in the second period of his life.

The planner would prefer not to split his homogeneous group of machines in this way if

$$(1-\gamma)\left[(1-\beta)c - 1 + F(i_p+1,N)\right] - \beta\left[F(1,N) - F(1,\gamma N)\right] < 0. \quad (17)$$

In particular, a sufficient condition for (17) to be satisfied is to have $F_{22} > 0$, which we will assume throughout[16]. The effect of this assumption on the behavior of the planner is clear: if the positive network externalities grow fast enough with the size of the user base, they will outweigh any incentive to delay the adoption time on a fraction of the machines used, since the planner will prefer to have a homogeneous group given the convexity of $F(.,n)$.

However, this is just one of the many types of alternative policies available to the planner. The following proposition shows that the intuition of the previous example goes through for the general case, and so the constrained solution is also the unconstrained solution for the planner when the payoff function is (weakly) convex in the number of agents using the same machine.

Proposition 1 *The constrained solution to the planner's problem is also the solution to the unconstrained problem if $F_{22} \geq 0$, that is, starting from a group of machines with the same vintage, the optimal policy is to follow synchronized adoption cycles of length $i_p(N)$.*

Proof. See Appendix.

4 Synchronized Equilibrium

From the set of open-loop equilibria of this game, equilibrium outcomes that involve synchronized adoption cycles are of particular importance for the objectives of this paper. In this section, we turn attention to equilibria where agents decide to synchronize their actions in common adoption dates. In particular, we will define a **synchronized** equilibrium as an open-loop equilibrium where agents coordinate their adoption dates, that is $s_j(t) = s(t)$ for all j.

[16] Notice that in (17) the left hand side is equal to zero for $\gamma = 1$, (since there is no one delaying adoption) and negative for $\gamma = 0$, since that is equivalent to synchronizing everyone into a $i_p(N)+1$ threshold, which is suboptimal, by the definition of the optimal policy $i_p(N)$.

Therefore, to prove that (17) is violated for any $\gamma \in [0,1]$ it is enough to show that the left side is (weakly) convex in γ. That condition is equivalent to $\beta N F_{22}(1,\gamma N) \geq 0$ which is equivalent to the claim since both β and N are positive.

Define a **periodic synchronized equilibrium** of length i —an i-cycle equilibrium—as a synchronized equilibrium where all agents switch to the same new machine every i periods, that is, for some integers i, k, and l,

$$s_j(t) = s(t) = \begin{cases} 0 & \text{for } t = ki + l \\ s(t-1)+1 & \text{otherwise} \end{cases} \quad (18)$$

The following proposition shows that when the effect of the externality is low enough, there exist (possibly multiple) periodic synchronized equilibria. Moreover, we can always assure in this case that the planner's solution is one of the equilibria in this case, a result that parallels Cooper and Haltiwanger (1992).

Proposition 2 ***(Existence of an efficient synchronized equilibrium for low N)*** *For $N < n_l$ an adoption cycle of length $i_p(N)$ constitutes an equilibrium [The planner's policy is an equilibrium policy].*

Proof. The proof is based on the planner being able to replicate any deviation of a single agent and, whenever the payoff function is always decreasing in the age of the machine ($N < n_l$), get a payoff that is no lower than that obtained by the deviator.

In particular, if the deviator decides to innovate before age $i_p(N)$, then he can get a machine of age 0 and a payoff that period of $F(0,0)$, something that is also attainable by the planner by innovating that period and getting a payoff $F(0,N) = F(0,0)$. Thereafter, the planner can follow the deviator's policy and get a payoff that is weakly higher each period (strictly higher if the deviator does not keep switching vintages every period).

If the deviator decides to *delay* adoption for t periods he can then rejoin the rest of agents by innovating to their vintage and get a payoff of $F(t,N)$ at the moment he rejoins. The planner can imitate the same policy, but of course cannot get a machine that has been used in the previous period by N users as the deviator does. The best the planner can hope for after t periods is to get a brand new machine (age 0). Nevertheless he is able to get a payoff that is no lower than that of the deviator, since for $n < n_l$, $F(0,N) > F(x,N)$ for all x.

Thus any deviation cannot improve over the payoff of synchronizing over the planner's cycle. But since the solution to the planner's problem already gives the maximum payoff attainable, then no such deviation can be profitable. So the result follows. ■

Notice that the condition $N < n_l$ used in the proposition is crucial because if the group size is bigger than n_l, then an agent may have the opportunity of delaying adoption and then adopt the vintage everyone in the group has been using for t periods, with $t \in \left\{1, 2, \ldots \hat{t}_n\right\}$[17] so that $F(t,N) > F(0,N) = 1$. The planner may not get a higher payoff by following the same policy —he cannot follow the deviator's policy and get more than $F(0,N) = 1$ at the time of innovation (that is, after t periods), since for the planner there is no vintage that has been used the previous period, so he only has a new vintage and unused old vintages to choose from. The problem for the planner is so serious that as it is shown below, when the group size is big enough ($N > n_l$) the planner's policy does not even constitute an equilibrium policy for low switching costs.

Although the previous proposition gives us an idea of one of the decentralized outcomes of this model, it is not by far a good description of all the possible outcomes. In particular, one of the questions left open by the previous result is the possibility of multiplicity of equilibria, something that seems a recurrent feature in dynamic games with positive payoff externalities.[18]

The following lemmas characterize the two forces underlying the individual decisions vis-à-vis the behavior of the rest of agents: on the one hand, an agent would have an incentive to delay adoption of a new machine since that defers incurring the fixed cost c, and more importantly, it allows him a "free ride" as other agents build up the externality associated with a particular vintage (the increasing part of $F(x,.)$). Preventing that kind of deviations gives a lower bound for the possible age thresholds that can constitute a synchronized equilibrium. On the other extreme, for sufficiently low adoption cost c, agents cannot wait indefinitely to adopt a new machine, since at some point the "stand alone" payoff of a new machine would outweigh the loss of the spillovers from leaving an old machine used by everyone. That gives an upper bound on any equilibrium threshold.

Lemma 1 ***(Sufficient conditions for a synchronized equilibrium threshold)*** *Suppose $N > n_l$, and all agents are following a synchronized i-cycle adoption policy. Then no agent would have an incentive to follow a different adoption policy if the following conditions are satisfied*

$$F(0,N) - F(i+1,N) > (1-\beta)\, c \tag{19}$$

$$F(0,N) - F(i+1,N) + \sum_{x=1}^{\hat{t}_N} \beta^x \left[F(x,N) - F(x+i+1,0)\right] > c \tag{20}$$

$$\underset{t\in\{1,2,\ldots i-1\}}{Max} \left\{ \begin{array}{c} \sum_{x=0}^{i-t} \beta^x \left[F(x,0) - F(t+x,N)\right] \\ +\beta^{i-t+1} \sum_{x=0}^{\tau(t)} \beta^x \left[F(i-t+x+1,0) - F(x,N)\right] \\ -c(1-\beta^{i-t+1} + \beta^{i-t+\tau(t)+1}) \end{array} \right\} < 0 \tag{21}$$

[17] Recall that $\hat{t}_N$ is defined by the maximum value of t satisfying $F(0,N) \leq F(t,N)$. This is the oldest age of the machine that gives an agent at least the same payoff as a new machine. Note that for $N < n_l$, $\hat{t}_N = 0$.

[18] See for example Krugman(1991) and Matsuyama(1991), and closer to this model, the results in Cooper and Haltiwanger (1992) and Shleifer(1986)

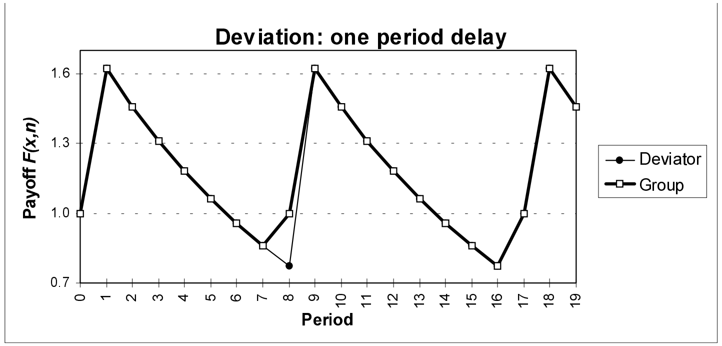

Figure 3: Deviation in equation (19)

where

$$\tau(t) = \left\{ \begin{array}{ll} 0 & \mathit{if}\ F(0,N) - F(i-t,0) > (1-\beta)c \\ Max\{\tau \leq t : F(\tau,N) - F(i-t+\tau,0) \leq (1-\beta)c\} & \mathit{otherwise} \end{array} \right.$$

Proof. See Appendix.

Notice that equation (19) is also a necessary condition to prevent delay by a single agent, since it represents the difference in payoff from one possible deviation by an individual agent, as represented in figure 3. In this case, an agent waits for one period after everyone else decides to adopt and then "rejoins the group" by adopting the same vintage they adopted the previous period (which now has age 1). However, equation (20) does not represent the net payoff of any feasible deviation —although this agent delays adoption for $\hat{t}_N$ periods, condition (20) does not include the cost of adoption of the group's technology at time $\hat{t}_N$—, so this is not a necessary condition to prevent delay.

Condition (21) has the opposite interpretation: it is required to prevent an agent from adopting a machine before age i (the group's policy). It thus gives an upper bound on the values of i that can constitute a (synchronized) equilibrium threshold in the decentralized solution. It also forms part of any necessary condition to prevent early adoptions since it represents the payoff of a possible deviation (early adoption) by a single agent, as in figure 4[19].

[19] The conditions of the lemma are similar in spirit to those in Shleifer (1986), although here agents are not restricted on how far they can wait to use their option to innovate. In Shleifer, agents can only wait up to one implementation cycle and still be able to grab the gains from implementing their new invention.

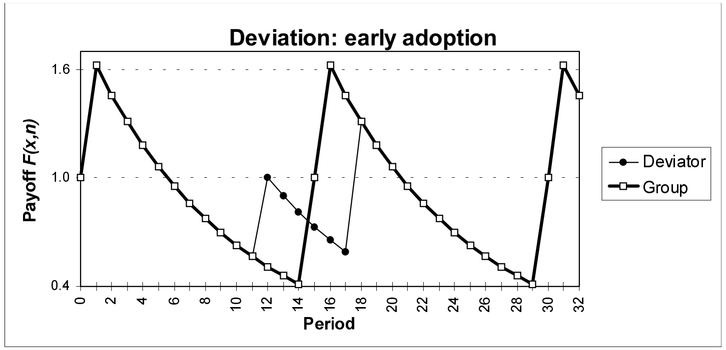

Figure 4: Deviation in equation (21)

Using this lemma, we can compare the conditions for a synchronized equilibrium with the planner's policy derived in the previous section. Although we showed that the planner's solution can be sustained as an equilibrium outcome for a user base N lower than n_l, the opposite is true when the user base is bigger than that level: *every* synchronized equilibrium involves too much delay, so agents coordinate on adopting new vintages with a frequency that is lower than optimal.

Proposition 3 ***(Inefficient delay for a big user base and low*** c***)*** *For any size of the group* $N > n_l$*, there exists a level of adoption cost* $c_N > 0$ *such that if* $c < c_N$ *then any synchronized equilibrium is inefficient:* $i > i_p(N)$.

Proof. Note that condition (19) in effect establishes that any synchronized equilibrium threshold i must be greater than $\hat{t}_N$. However, for $c = 0$ and $N > n_l$ the planner's solution $i_p(N)$ must be no greater than $\hat{t}_N$ —since at $i_p(N) = \hat{t}_N + 1$ the right side of (15) is bigger than 0. By continuity around $c = 0$, for small values of c, the planner's solution does not constitute an equilibrium. ■

The preceding proposition shows that for a group size big enough, *any* synchronized equilibrium is inefficient. Note that when the group size is bigger than n_l, an agent has an incentive to delay adoption with respect to a prescribed synchronized rule because of two reasons. The first is that delaying adoption postpones incurring the switching cost c. Second, an agent has an incentive to delay adoption and wait for the rest of agents to "build up" the externality associated with a particular vintage —the increasing part of $F(x, \cdot)$—, a "free ride" that the planner cannot take

advantage of. This last reason is the most important, and the only one that drives the result of proposition 3, since it also holds for $c = 0$.

The threshold n_l determines how big the group needs to be in order to generate this inefficiency. For a user base bigger than n_l the payoff function becomes increasing in the age of the machine during the first period (see figure 1); this allows an individual agent to delay adoption for a few periods (less than $\hat{t}_N$) respect to the planner's policy and rejoin the group afterwards. Doing that gives that agent a higher payoff than following the planner's solution. Obviously, if the planner replicates this policy, he cannot "free ride" on himself, as we argued in proposition 1.

One particularly striking result is the possibility of inefficient collapse of investment in new vintages. Suppose that $N > n_l$ and $\sum_{t=0}^{\infty} \beta^t F(t, N) > c \geq \frac{F(0,N)}{(1-\beta)} = \sum_{t=0}^{\infty} \beta^t F(0, N)$. Note that in this case[20] condition (19) would be violated, which means that for any adoption threshold i that may be proposed as an equilibrium adoption date, an individual agent would always have an incentive to delay adoption for one period and then rejoin the others when they have a machine of age 1 —the deviation in figure 3. However, for these values of c, the planner's solution *does* involve periodic innovations[21]. Thus, in this case, there is **inefficient stagnation**: any synchronized equilibrium involves no innovation ever, even though it is socially optimal to have periodic innovations.

Although an important result, the previous proposition only asserts the properties of the equilibria for a user base high enough. The next proposition tackles the problem of existence of such equilibria, at least for high values of the user base N.

Proposition 4 ***(Existence of a synchronized equilibrium for high*** N***)*** *There exists a number* n_h *such that for all* $N > n_h$ *a synchronized equilibrium exists.*

Proof. By analyzing (19) and (20) we can show that in both cases the left side of the inequalities is increasing in i. Therefore, we can establish the existence of a number i_1 such that (19) is satisfied if and only if $i \geq i_1$. Equivalently, there exists a number i_2 such that $i \geq i_2$ is a necessary and sufficient condition for (20) to be satisfied.

On the other hand, one can show that the left side of (21) is increasing in i. To see this, denote by $t^*(i)$ the optimal t associated with a particular threshold i in the maximization on the left side of the inequality (21). Now increase i by one unit to $i+1$ and take $t^*(i)+1$ as the choice for t. Note first that $(i+1)-(t^*(i)+1) = i-t^*(i)$ and so $\tau(t;i) = \tau(t^*(i)+1; i+1)$. Therefore, all the terms involving $i - t$ and $\tau(t)$ remain unchanged. That leaves only the first term $\sum_{x=o}^{i-t} \beta^x \left[F(x, 0) - F(t + x, N)\right]$

[20] Note that this condition will be will be fulfilled for $c > \frac{1}{1-\beta}$ and high values of N.

[21] Equation (15) would be satisfied for some arbitrarily large but finite i, as $\lim_{i\to\infty} F(i, N) = 0$ for all N

which is increasing in t[22]. Since $t^*(i)+1$ is an arbitrary choice of maximizer to the problem in (21) with $i+1$, then the maximum value attained in the left side of (21) must be strictly higher with $i+1$ than with i.

Given that the left side of (21) is increasing in i, then there exists a number i_3 (possibly infinite) such that (21) is satisfied if and only if $i \leq i_3$. Therefore, an equilibrium exists if there is a number i satisfying

$$Max\{i_1, i_2\} \leq i \leq i_3$$

We can first show that $i_1 < i_3$. To see this, analyze the maximization problem on the left side of (21) when $i = i_1$. Since i_1 is defined as the minimum value of i satisfying (19) then we have that for all $x \in \{0, 1, \ldots i_1 - t\}$ $F(x,0) - F(t+x,N) \leq F(0,N) - F(i_1,N) < (1-\beta)c$, which means that the left side of (21) becomes

$$\begin{aligned}
& \underset{t\in\{1,2,\ldots i_1-1\}}{Max} \left\{ \begin{array}{c} \sum_{x=0}^{i_1-t} \beta^x \left[F(x,0) - F(t+x,N)\right] \\ +\beta^{i_1-t+1} \sum_{x=0}^{\tau(t)} \beta^x \left[F(i_1-t+x+1,0) - F(x,N)\right] \\ -c(1-\beta^{i_1-t+1} + \beta^{i_1-t+\tau(t)+1}) \end{array} \right\} \\
< & \underset{t\in\{1,2,\ldots i_1-1\}}{Max} \left\{ \begin{array}{c} \sum_{x=0}^{i_1-t} \beta^x \left[(1-\beta)c\right] + \beta^{i_1-t+1} \sum_{x=0}^{\tau(t)} \beta^x \left[F(i_1-t+x+1,0) - F(x,N)\right] \\ -c(1-\beta^{i_1-t+1}) - c\beta^{i_1-t+\tau(t)+1} \end{array} \right\} \\
= & \underset{t\in\{1,2,\ldots i_1-1\}}{Max} \left\{ \begin{array}{c} \frac{(1-\beta^{i_1-t+1})}{(1-\beta)} \left[(1-\beta)c\right] + \beta^{i_1-t+1} \sum_{x=0}^{\tau(t)} \beta^x \left[F(i_1-t+x+1,0) - F(x,N)\right] \\ -c(1-\beta^{i_1-t+1}) - c\beta^{i_1-t+\tau(t)+1} \end{array} \right\} \\
= & \underset{t\in\{1,2,\ldots i_1-1\}}{Max} \left\{ \beta^{i_1-t+1} \sum_{x=0}^{\tau(t)} \beta^x \left[F(i_1-t+x+1,0) - F(x,N)\right] - c\beta^{i_1-t+\tau(t)+1} \right\} < 0
\end{aligned}$$

where the last inequality comes from realizing that all the terms in the summation are negative. Therefore, since for $i = i_1$ inequality (21) is satisfied, and the left side is increasing in i we can conclude that $i_1 \leq i_3$.

The last part of the proof is to show that for N big enough, $i_1 > i_2$. To prove this notice that if we set $i = i_1$ in inequality (20), then the term $F(0,N) - F(i_1+1,N)$ is decreasing in N, with a lower bound equal to $(1-\beta)c$. However, the second term $\sum_{x=1}^{\hat{t}_N} \beta^x \left[F(x,N) - F(x+i+1,0)\right]$ is increasing in N and unbounded (Recall that $\hat{t}_N$ is an increasing function of N). Therefore, there must exist a number n_h for which $i_2 < i_1$ and thus $Max\{i_1, i_2\} \leq i \leq i_3$. Since $F_{12} < 0$ that also holds for any $N > n_h$. ■

One problem with the above proposition is that in general, it cannot rule out that $n_l < n_h$ and so the existence of a synchronized equilibrium for intermediate values of n cannot be guaranteed.

[22] Recall that the difference $i - t$ is kept constant (by the increase in i) and so only the term $F(t+x,N)$ decreases as t increases. The number of terms in the summation is constant.

It is also worth noting that the proof of this lemma does not rely on checking that the planner's solution constitutes an equilibrium, as proposition 2, precisely because when N is big enough, the planner's policy cannot be sustained in equilibrium, as we have already shown.

Although this paper is concerned with the behavior of individual units that are forward-looking, it is interesting to mention the case of zero adoption costs. With no switching costs, agents always switch to the vintage with the highest per-period payoff, irrespective of the future evolution of the distribution of vintages in use. However, this case is illustrative of the role of lagging externalities and a non-monotonic payoff function in the results drawn so far. In particular, this shows that the "cost" of building up the externality associated with a particular vintage and the incentives it gives for individual free riders are the main driving forces behind the appearance of inefficient delay in the model.

Proposition 5 ***(Uniqueness of the (Inefficient) Synchronized Equilibrium for*** $c = 0$***)*** *If* $c = 0$*, there exists a unique steady state synchronized equilibrium for any group size* n*. Moreover, in this equilibrium agents follow cycles of length* $\hat{t}_N$ *and is thus inefficient for* $N > n_l$*.*

Proof. The proof relies on the fact that for $c = 0$, agents always switch to the vintage with the highest current payoff, regardless of the future path of payoffs associated with any vintage. Thus, any equilibrium policy can neither call for an agent to keep his machine past age $\hat{t}_N$ —where the payoff is lower than that of a new machine— nor to innovate before $\hat{t}_N + 1$ —since he would be switching to a machine with lower current payoff than the one he is using. Thus the only possible candidate for a synchronized equilibrium is an innovation cycle of length $\hat{t}_N$, which satisfies the sufficient conditions specified in lemma 1. The inefficiency of this unique equilibrium is a direct consequence of proposition 3. ■

This result also shows that —as would be expected—, the irreversibility associated with the fixed cost c is the source of the multiplicity of equilibria in this model, since agents have to take into account the expected evolution of the distribution of machines over vintages when deciding their optimal policies.

5 Staggered Equilibrium

Up to this point we have focused on describing a particular steady state, namely, one in which all agents start with machines of the same vintage. The results we have drawn are, first, that perfect synchronization is an equilibrium in this model, one that transmits the non-convexities present at the individual level into the aggregate (macroeconomic) time series. Second, the periodicity of actions is reduced inefficiently in equilibrium, so the interaction of individual agents not only fails to smooth out the nonconvexities found at the individual level, but it also amplifies them, and leads to too much delay in equilibrium.

In this section we investigate the other extreme aggregate behavior in the steady state: whether we can still sustain complete smoothing of individual lumpy actions in the face of positive payoff externalities. As opposed to the previous case in which we "start" the world with a distribution of machines over vintages concentrated on a single vintage, here we will ask whether a non-degenerate distribution can survive in the steady state.

Let n_{xt} denote the mass of users of a machine of age x at time t, with $x = \{0, 1, 2, \ldots\}$. A distribution of machines over vintages is then given by $n_t = \{n_{0t}, n_{1t}, n_{2t}, \ldots\}$, with the obvious requirement that $n_{xt} \geq 0$ and $\sum_{x=0}^{\infty} n_{xt} = N$. We will refer to a **degenerate distribution** at time t when $n_{it} = N$ for some i, which is the case analyzed when we dealt with synchronized equilibria. Accordingly, a **non degenerate** distribution corresponds to having $0 < n_{it} < N$ for some i, which immediately implies that at least two vintages are being used simultaneously at time t.

Define a **homogeneous** distribution as a non degenerate distribution with $n_{it} \in \left\{0, \frac{N}{k}\right\}$, where k denotes the number of vintages with a positive measure of users at time t. In words, a homogeneous distribution implies subgroups of machines of the *same* size using different vintages at time t. Note that this definition imposes no restriction on the location of the k vintages being used.

Define a **periodic staggered equilibrium** as an open loop equilibrium where $s_j(t) \neq s_{j'}(t)$ for some $j \neq j'$ (and there is a positive measure of agents identified with index j and j' respectively), but both s_j and $s_{j'}$ have the *same* period length i. (as defined in (18)). Thus, in any staggered equilibrium and at any time t, we have at least two different vintages being used simultaneously by groups of agents of positive measure.

The following proposition shows that, starting from a homogeneous distribution of machines and for a switching cost high enough, there exists multiple staggered equilibria in this model, which in fact says that another possible steady state for this model implies groups of agents leapfrogging each other over the technology adoption cycle.

Proposition 6 ***(Existence of staggered equilibria)*** *Suppose we start with a homogeneous distribution of machines over vintages. Then there exists a number $c_l > 0$ such that, for $c > c_l$, any group size N, and any $k \geq \frac{N}{n_l}$ there exists (many) staggered equilibria in which k groups of equal*

size $\frac{N}{k}$ follow a periodic cycle of length $i_p(\frac{N}{k})$.

Proof. By lemma 2 we know that for a group of size N smaller than n_l there exists an equilibrium in which the agents follow the planner's policy (innovation every $i_p(N)$ periods). Suppose all agents in each of the k subgroups of size $\frac{N}{k}$ update their machines in cycles of length $i_p(\frac{N}{k})$. Since it is an equilibrium to follow that policy when there is just one group of that size, then the only possible profitable deviation for an individual agent is to switch groups.

Suppose at time t an agent is originally in a subgroup (call it group A) and decides to switch to subgroup B. Since by hypothesis both groups have the same size below n_l —thus the requirement of $k \geq \frac{N}{n_l}$—, then the per-period payoff is everywhere decreasing in the age of the machine, and so a planner can always follow the deviator's policy and get an average payoff per machine that is no lower than that of the deviator.[23] Since agents were already following a cycle of length $i_p(\frac{N}{k})$ —the planner's solution for a single group of size $\frac{N}{k}$— then that kind of deviation is not profitable. Thus there is no profitable deviation from the proposed staggered path, and so staggered strategies of this type constitute an equilibrium .

Since the proof of this proposition does not depend in any way on the distribution of the k groups over the $i_p(\frac{N}{k})$ possible vintages, it follows that there are multiple staggered equilibria with k subgroups. Moreover, the number of possible subgroups is only restricted to be higher than $\frac{N}{n_l}$ and obviously lower than $i_p(\frac{N}{k})$ which increases the multiplicity in that dimension as well. ■

Note that this proposition, since it deals with homogeneous distributions in the steady state, allows for any size of the subgroups (between $\frac{N}{i_p(\frac{N}{k})}$ and n_l) and any distribution of the groups of size $\frac{N}{k}$ across vintages. In particular, note that the subgroups can use machines of very close vintages and still there would be no incentive to switch groups, even when that may imply just adopting a new machine one period earlier or later than prescribed. The fact that they are following the planner's policy for each individual group does not mean that a planner would not prefer to merge some of the groups together —especially in cases where the vintages used are very close[24]—, since the gains in the increased externality would outweigh paying the cost of adoption earlier than

[23] See the proof of proposition 2 and the discussion that follows it. The kinds of deviations attainable by an individual agent in the two cases are essentially the same, and the purpose of the assumption $k \geq \frac{N}{n_l}$ is also equivalent: only with a subgroup size that is below n_l we can assure that the payoff function F is decreasing in the age of the machine, and so, by following the same policy of the deviator, the planner can always attain a per-period payoff that is no lower that that of the former.

[24] Or very far away: note that in all this analysis, everything must be taken modulo $i_p(.)$. Thus, two subgroups that follow the same adoption threshold i_p and at time t use technologies with an age difference of $i_p - 1$, will have technologies with just one year difference next period, since it is the turn of the group having an older technology to adopt a new vintage.

prescribed by the equilibrium policy.

Note that when the switching cost c is close enough to zero, a staggered equilibrium of the kind mentioned in the proposition does not exist for $N > n_l$. The reason is that such a staggered equilibrium would call for all agents updating their machines to a new technology every period[25], effectively synchronizing them into a single vintage. But then, given that $F(1, N) > F(0, N) = 1$ for $N > n_l$, an agent would prefer not to adopt continuously and hold his machine for more than one period.

The fact that we can construct a staggered equilibrium of this sort for subgroups close enough in the age of their vintages would, in principle, provide a complete smoothing of the individual non-convexities, giving a time series of investment without cycles[26]. However, note that the proposition relies heavily in the fact that the groups are *exactly* of the same size and thus the planner of a subgroup can always replicate the payoff obtained by an agent switching groups. That of course is not true with subgroups of slightly different size and with small switching costs, and thus suggests that a staggered equilibrium of this sort is not stable to a small perturbation in the size of the k subgroups, especially when the vintages are very close in age (that is, when the time period is small enough).

One way in which we can see that synchronization in this context is the "natural" outcome is to consider the particular case of zero costs[27]. In the next section, we show that for that particular case, the unique (globally stable) steady state involves synchronization, which reinforces our case. The previous discussion also shows that staggered equilibria (that could eventually smooth out the lumpiness at the individual level) could disappear if we introduce small perturbations in the distribution of machines across vintages. Another way to introduce perturbations would be to consider that the evolution of the "state of the art" technology is still exogenous, but uncertain. In section seven we show that even small aggregate shocks move the system away from a steady state with complete smoothing.

[25] Recall that the size of each subgroup is smaller than n_l, and so the payoff function $F(x, n)$ is decreasing in the age of the machine. Then, when the adoption cost is zero, there cannot be any equilibrium in which an agent receives a payoff lower than $F(0, n) = 1$ in a particular period, since there is always the option of adopting a new machine alone in that period and get $F(0, 0) = 1$.

[26] A similar result in the context of the rigidity of prices and the effect on macro variables is found in Caplin and Spulber (1987).

[27] Recall our previous discussion on the relative roles of the switching costs and the lag in the network externalities. The important assumption to obtain the results about inefficient delay in this model come from the introduction of lags in the network externalities, whereas switching costs force agents to be forward looking.

6 Convergence

In the previous two sections we have analyzed two possible candidates for steady state equilibrium strategies. In the first case, we showed that starting from a degenerate distribution, synchronization can be sustained in equilibrium. On the other hand, starting from a homogeneous distribution, staggering of adoption dates and eventually a complete smoothing of the aggregate time series is possible, as shown on the previous section, although we argued that this last case is just a knife-edge situation.

To further reinforce the intuitive idea that synchronization should be the natural outcome in the presence of network externalities, we want to consider the issue of stability, that is, starting from any arbitrary distribution of machines over vintages, which of the two steady states is the one that arises in the long run. The general answer to this question is not straightforward, and is extremely difficult to answer for the general case. However, the particular case for zero adoption costs yield a clear cut answer that sheds light into the more general question.

As before, let n_{xt} denote the mass of machines of vintage x being used at time t and $n_t = \{n_{0t}, n_{1t}, n_{2t}, \ldots\}$ denote the distribution of machines over vintages at time t. For any initial distribution of machines over vintages n_t, there exists an associated distribution of *payoffs* over vintages $\pi_t = \{\pi_{0t}, \pi_{1t}, \pi_{2t}, \ldots\}$, where $\pi_{it} = F(i, n_{i-1,t-1})$ and $\pi_{0t} = 1$ for all t. Let X_t be the set of vintages that provide the maximum payoff at time t, that is

$$X_t = \left\{ \arg\max_i \{\pi_{it}\} \right\}$$

and let ξ_t the number of elements in X_t. The following proposition shows that in the particular case of zero switching costs, the unique globally stable steady state involves synchronized cycles of length $\hat{t}_n$, so there is no smoothing by aggregation. Moreover, for $n > n_l$, proposition 5 applies so the steady state is inefficient.

Proposition 7 ***(Global Stability of the unique Synchronized Equilibrium for*** $c = 0$***)*** *For a switching cost* $c = 0$*, and for any initial distribution of machines over vintages* n_0*, all agents synchronize into a cycle of length* $\hat{t}_N$ *after at most* $\hat{t}_N$ *periods.*

Proof. For $c = 0$, any equilibrium path involves agents switching immediately to the vintage with the current highest payoff, regardless of their expectations about the future path of the distribution of machines over vintages (note that the short-run payoff function adjusts continuously to changes in the size of the cohort). In the absence of any tie in the initial profitability of two or more vintages

—i.e. when $\xi_0 = 1$—, this means that all agents will concentrate on the vintage with the highest initial payoff and follow a $\hat{t}_N$ cycle thereafter by proposition 5, so in this case agents coordinate into a $\hat{t}_N$ cycle after just one period.

If there are more than one vintage giving the maximum payoff at the initial date ($\xi_0 \geq 2$), then any equilibrium policy involves agents switching immediately (at time 0) to vintages in the set X_0. The distribution of payoffs over vintages after this initial switch π_1 can only be a degenerate distribution —if all agents coordinated to switch into the same vintage in the set X_0—, or a homogeneous distribution with only vintages in X_0 having a positive mass of users. Any other distribution π_1 could not be an equilibrium since there would be an incentive for an individual agent to switch from a vintage with a lower payoff to a vintage with the maximum payoff, because of zero switching costs.

If every agent coordinates in period 1 to use the same vintage, then they will follow a synchronized cycle afterwards. If two or more cohorts with the same payoff in period 1 survive, then we know that each cohort cannot keep a machine older than $\hat{t}_n$ afterwards, since the payoff of a new machine would then be higher and there are no switching costs. Suppose these cohorts are able to keep their payoffs equalized in all the periods after the switching at time 0 (if not, they would switch to the vintage that develops a higher payoff, and we are again in the case of complete synchronization).[28] Then they will eventually arrive to a payoff equal or lower than 1 (the payoff of a new machine) at the same time. At that moment, all agents in all cohorts have to switch together to adopt a new machine. From that point onward, all agents have to follow a synchronized cycle of length $\hat{t}_N$. ■

7 Uncertainty and Aggregate Shocks

In this section we extend the framework developed in the basic model by analyzing the effect of introducing uncertainty into the model. We will assume that the evolution of the "state of the art" technology is not deterministic, but increases with random "jumps" in productivity. This type of setup seems more natural when analyzing the problem of consumers who face an uncertain prospect of new technologies coming into the market, since generally they have very limited information

[28] This ability to keep the same payoff for, say, l periods would suppose that, for example, in the case of two surviving cohorts of ages x_1 and x_2 and sizes n_1 and n_2 respectively, we would need

$$F(x_1, n_1) - F(x_1 + k, n_1) = F(x_2, n_2) - F(x_2 + k, n_2),\ \forall\ k \leq l$$

which violates our assumption $F_{12} < 0$. However, to show that this assumption is not crucial for the validity of the argument, we can assume for a moment that the condition above is satisfied.

about the capabilities of new models being developed by suppliers.[29]

As opposed to the deterministic case, the presence of aggregate shocks in the evolution of technological progress generates a tendency for further clustering of agents in a few vintages, thus breaking the "counterintuitive" staggered equilibrium found in the previous chapter. The intuition for this result is clear: as a shock generates a big increase in the productivity of the new model coming into the market, it may be more profitable for some agents to bring forward the time of adoption of a new model, since the extra burden of paying the switching cost in advance is more that compensated by the extra gains in productivity of the new model. This means that two groups that are using machines of very close vintages and that are also close to their (endogenous) scraping age will, in the face of a big productivity shock, decide to adopt the new model *at the same time*, thus collapsing two subgroups into a single one, and breaking a possible complete staggering configuration.

In order to model a stochastic improvement in the technological frontier, we will assume that existing vintages become "older" only as a new vintage appears. Moreover, the aging process may involve discrete jumps (of more than one period) on the age of all existing vintages. Recall that in the simpler model, the appearance of a new vintage (which occurred every period), meant that all existing vintages increased their index (age) in one unit. Thus, a machine with index i that had been used by n agents in the previous period had a productivity today equal to $F(i, n)$,and would have a productivity next period equal to $F(i + 1, n)$,provided there was no change in the number of users.

In the stochastic version of the model, we will consider that each period a new machine appears in the market, but the relative profitability of that machine with respect to all existing machines is unknown until the time of its appearance. Thus existing machines with index i last period will have an index equal to $i + l$ this period, with l a natural number bigger or equal to zero.[30]

Our main purpose in this section is to show that the introduction of aggregate uncertainty reinforces the tendency of agents to synchronize their adoption timing. In particular, we are interested in whether complete staggering can still be sustained as an equilibrium in an uncertain environment. For this, we develop first a very simple case, in which there is a one-and-for-all aggregate shock that increases technological progress, starting from a steady state of complete staggering.

[29] Another interpretation for uncertainty in the productivity of new machines is also the extent to which a support industry, or learning effects will be important for each model. However, we consider only the case of uncertainty in the intrinsic characteristics of the new model.

[30] Note that the case of $l = 0$ includes the case in which there is no innovation in this period, and thus, no new technologies appearing in the market.

Consider a steady state in which we have some degree of smoothing by aggregation by having k groups (of possibly different size) using different vintages at the same time. In period 0, and in the deterministic world of the previous sections, the group using the oldest vintage would update their machines to vintage 0, which just appeared in the market, and all other groups would have machines with an index increased by 1 unit. Suppose now that instead, we have a once-and-for-all jump in technological progress, and thus, at time 0, all machines that are not updated will have their index increased by l units, with $l \geq 2$[31]. We are interested in whether the complete staggered equilibrium can still be sustained after this shock. The following proposition shows that for any staggered equilibrium, the presence of aggregate shocks that are big enough will tend to reduce the number of vintages in use, by inducing the simultaneous adoption of a new vintage by at least two groups.

Proposition 8 *For any distribution θ corresponding to a staggered equilibrium with finite cycles ($i < \infty$) there exists a number l_θ, such that, if $l > l_\theta$, then an aggregate shock that increases the leading technology l steps will induce at least two of the subgroups with the oldest technologies to switch to the same vintage.*

Proof. See Appendix ■

A direct corollary of this proposition is then that the complete staggered equilibrium that would bring about complete smoothing in the aggregate time series of investment will not be sustained in the presence of an aggregate shock. This confirms the intuition that complete smoothing is a knife edge case in the deterministic case.

Proposition 8, however, does not specify the magnitude of the shock that would be necessary to break the complete staggering configuration that gives rise to a smooth aggregate time series. In order to tackle that problem we confine attention to one extreme: a particular case in which there are k groups of size $\frac{N}{k} \leq n_l$ using the first k vintages available (from vintage 0 to vintage $k-1$). In a deterministic setting the equilibrium strategy as a function of the vintage of the machine in use i would be

$$s_j(i) = s(i) = \begin{cases} 0 & \text{for } i = k \\ s(i-1)+1 & \text{otherwise} \end{cases}$$

Consider now an aggregate shock of size $l \geq 2$. As the shock occurs, agents with vintage $k-1$ are left with machines of vintage $k-1+l > k$ and agents with vintage $k-2$ are left with vintage $k-2+l \geq k$. Suppose that agents with vintage $k-1+l$ decide to update to technology 0, but not agents with vintage $k-2+l$, and consider the incentives of an agent who is in the latter group.

[31]The case $l = 1$ is the same as the deterministic rate of technological progress of the previous sections.

Suppose that complete staggering could be sustained in equilibrium even after the shock in period 0. That would require having all groups update to a new machine at the prescribed time before the appearance of a shock. In particular, it would mean that the group with vintage $k-2+l$ decides not to update to the new technology in period 0, but waits until next period to update to a new model.[32]

Take now one agent from the group who decides to stay with vintage $k-2+l$ in period 0. If that agent decides to switch to the new model (that has been adopted by $\frac{N}{k}$ users) in period 0 instead, he will bring forward the payment of the switching cost, and thus the strategy will have a cost of $(1-\beta)c$. On the other hand, the switch to the new vintage allows him to enjoy the benefits of a higher payoff in period 0 although he will have an older technology for the rest of the cycle (since he innovated earlier than his group). That represents a gain equal to

$$\left[1 - F(k-2+l, \frac{N}{k})\right] + \beta\left[V\left(1; \frac{N}{k}\right) - 1 - \beta V\left(1; \frac{N}{k}\right)\right]$$

where the first brackets enclose the difference in the first period in which the group switching is made, and the second bracket represents the difference from the second period onwards from staying in two groups of the same size using vintages with a difference of one period. Noting that we can equate $V(0; \frac{N}{k}) = 1 + \beta V(1; \frac{N}{k})$, an agent would then have an incentive to switch groups in the way described before if

$$-F(k-2+l, \frac{N}{k}) + \left[V\left(0; \frac{N}{k}\right) - \beta V\left(0; \frac{N}{k}\right)\right] - (1-\beta)c > 0$$

Using equation (16) we can replace the value of $V\left(0; \frac{N}{k}\right)$ and thus the condition for the complete staggering to be broken is equivalent to

$$-F(k-2+l, \frac{N}{k}) + (1-\beta)\frac{-c\beta^k + \sum_{x=0}^{k-1}\beta^x F(x, \frac{N}{k})}{1-\beta^k} - (1-\beta)c > 0 \qquad (22)$$

Note, however, that the left side of the previous inequality is equivalent to

$$\begin{aligned} & -F(k-2+l, \frac{N}{k}) + (1-\beta)\frac{-c + \sum_{x=0}^{k-1}\beta^x F(x, \frac{N}{k})}{1-\beta^k} \\ = & -c + \sum_{x=0}^{k-1}\beta^x F(x, \frac{N}{k}) - \sum_{x=0}^{k-1}\beta^x F(k-2+l, \frac{N}{k}) \\ = & -c + \sum_{x=0}^{k-1}\beta^x \left[F(x, \frac{N}{k}) - F(k-2+l, \frac{N}{k})\right] > 0 \qquad \forall l \geq 2 \end{aligned}$$

[32]More generally, it would imply that during the first k periods, all agents decide to use an adoption threshold age i equal to $k+1$.

where the inequality is a direct consequence of the properties of the planner's solution, as described in equation (15). Of course if one agent has an incentive to switch from the group that stays with the old technology to the group that adopts at time zero, then every other agent would want to make that switch. We would therefore have a collapse of (at least) the two group with the oldest technologies into a single one, thus generating a non-uniform distribution, and thus some non-smoothing in the next cycles. We can summarize the result in the following proposition

Proposition 9 *Starting from a steady state of complete staggering in which all the first k vintages are being used by groups of the same size, an aggregate shock of any size will force the merger of at least the two groups using the oldest technologies.*

These two propositions then show the robustness of the main results of the paper. The introduction of aggregate uncertainty generates a tendency for agents to synchronize their otherwise staggered decisions. That is, for big enough aggregate shocks, there will be a merging of two groups of agents that were using different vintages up to the time of the shock.

This merging in the face of an aggregate shock has two important implications. First, it shows that the complete staggering equilibrium that we could construct when the technological frontier evolves deterministically cannot be sustained in an uncertain environment, thus giving support to the idea that it was a knife edge case, not robust to exogenous perturbations. Second, the fact that there is always a shock big enough to merge two subgroup using different vintages, ought to give credit to the idea that the system exhibits a natural tendency for synchronization of actions. What this means at the aggregate level is that we will not see too much smoothing by aggregation (in fact, no smoothing if there is complete synchronization), and thus the microeconomic non-convexities get amplified at the macro level. It should also be clear that the introduction of aggregate shocks does not eliminate the tendency of agents to delay their adoption time respect to the socially optimal outcome.

8 Concluding Remarks

We have shown a model that combines a lumpy process at the individual level with a particular timing of network externalities. The results of the model imply that in this economy there will (generally) not be smoothing of individual decisions by aggregation, in accordance with some empirical and theoretical findings in the literature[33]. This tendency for synchronization arises even without aggregate shocks that act as a coordination force, as in Bertola and Caballero (1990). However, the introduction of lagging externalities into the model means that when the extent of the spillovers is significant, *all* equilibria in the model will exhibit inefficient *delay*, when previous literature could not rule out that agents could coordinate on an equilibrium with efficient cycles, as in Shleifer (1986) and Cooper and Haltiwanger (1992). Thus, in this model —and for externalities big enough— aggregation *inefficiently* increases the volatility of the time series of investment, instead of smoothing out individual non-convexities.

[33] See for example the work of Cooper, Haltiwanger and Power (1995) and Bertola and Caballero (1990) using U.S. data at the micro level for replacement processes that would fit into the kind of non-convex problem analyzed here. Their models, however, do not include any network externality affecting agents' payoffs.

There are still some open questions in this analysis. The first one is the complete characterization of the planner's solution in the presence of network externalities when the user base is big enough. Starting from a more general distribution of machines over vintages, a planner would in principle "settle" into having a non degenerate distribution, with two or more cohorts coexisting at any time. However, the influence of the (lagging) network externalities would force those two cohorts to be regularly spaced across the support of vintages available. The same pattern would be expected of the decentralized solution, thus providing support to the non-smoothing of the micro non-convexities in this model. Once again, the free riding problems that we encounter when characterizing the decentralized solution would be present also in this context, so we should expect inefficient delay as a robust phenomenon.

One particular point that could be criticized in this model as a reflection of some macroeconomic issue is the idea of network externalities introduced in vintages of machines. It could be argued that this model can be a model of the behavior of *one* sector of the economy, where this kind of externalities could be more plausible, as opposed to the interaction among different sectors. However, the same kind of issues encountered here would be found in a model where the positive externality works not at the level of payoffs, but through the switching cost. One could think of a model where the cost of innovation is lower the higher the number of people innovating at the same time, because of scale effects in the production of capital goods, or pecuniary externalities from investment, as in Shleifer (1986). If the positive externalities take some time to set in as in this model, the incentives to wait for others to invest and generate those spillovers appear again, and so inefficient delay would arise in equilibrium.

Another avenue of research necessary to understand the effect of network externalities in the smoothing of individual lumpiness would be to consider a stochastic environment, where agents receive common and idiosyncratic shocks (in productivity, or in the cost of switching to a different machine). The interaction of a replacement problem with the evolution of the distribution of agents over vintages has been studied by Cooper, Haltiwanger and Power (1995), in the context of isolated agents. The introduction of network externalities in that setting could be an interesting step forward in the study of economic fluctuations.

A Appendix: Proofs

A.1 Proof of Proposition 1

Consider a planner with machines of age $i_p(N)$ and take that as period 0. The optimal policy for a constrained planner is to innovate to new machines (of age 0) at that moment. Suppose instead that this planner updates only a fraction γ of his machines and keeps the remaining fraction $(1-\gamma)$ at the same vintage (no innovation).

Suppose the two subgroups eventually share the same vintage again in the future, and call t the first period in which that happens. Without loss of generality suppose also that following t the two groups remain synchronized, since if it is not profitable to split the groups once, it will never be profitable to split them again. The argument sketched in the text essentially showed that if $t = 1$ then the planner is better off by keeping an homogeneous group of machines when $F_{22} \geq 0$.

Case I: $1 < t < i_p(n)$

If $1 < t < i_p(N)$, that is, if the two groups rejoin before the completion of another cycle for the constrained planner, then the payoff from this strategy is given by

$$\gamma \left[-c + 1 + \sum_{x=1}^{t-1} \beta^x F(x, \gamma N) \right] +$$
$$+(1-\gamma) \left[F(i_p + 1, N) + \sum_{x=1}^{t-2} \beta^x F\left(i_p + 1 + x, (1-\gamma)N\right) + \beta^{t-1} F(t-1, \gamma N) - \beta^{t-1} c \right] +$$
$$+\beta^t V(t; N)$$

where the first line shows, respectively, the gains from the group that innovates at time zero, paying the cost c and having an externality given by a size γN during t periods, and the gains from the group of size $1-\gamma$ which maintains the externality of the whole group for the first period, then for the next $t-1$ periods sees the effect of that externality reduced to $(1-\gamma)N$ and then pays the adoption cost c to rejoin the first group when they use a vintage of age $t-1$. Note in this last case the effect of the lagging externality, since in the first period after the two subgroups are rejoined they still get the externality associated with that vintage in the previous period, that is, a user base of size γN.

To determine whether this split of machines is more profitable than a synchronized policy, we can compare the payoff over these t periods to the convex combination of two planners with adoption thresholds coincident with the ones used by the two subgroups (i_p and $i_p = t$) ,and using weights equal to the fraction of machines diverted to each subgroup. This convex combination of the two planners' payoffs would be given by

$$\gamma\left[-c+1+\sum_{x=1}^{t-1}\beta^x F(x,N)\right]+$$
$$+(1-\gamma)\left[F(i_p+1,N)+\sum_{x=1}^{t-2}\beta^x F\left(i_p+1+x,N\right)+\beta^{t-1}F(0,N)-\beta^{t-1}c\right]+$$
$$+\beta^t\left[\gamma V(t;N)+(1-\gamma)V(1;N)\right]$$

Note in particular that the planner following the subgroup that delays, may not be able to achieve the same payoff after adoption, since in that case, the subgroup is taking advantage of a vintage that has been used in the previous period, something the planner with the delayed adoption date does not have available.

The difference between the two payoffs is given by

$$\begin{aligned}
& \gamma\left[\sum_{x=1}^{t-1}\beta^x\left(F(x,N)-F(x,\gamma N)\right)\right]+ \\
& +(1-\gamma)\left[\sum_{x=1}^{t-2}\beta^x\left(F\left(i_p+1+x,N\right)-F\left(i_p+1+x,(1-\gamma)N\right)\right)\right]+ \\
& +\beta^{t-1}(1-\gamma)\left[F(0,N)-F(t-1,\gamma N)\right]+\beta^t\left[\gamma V(t;N)+(1-\gamma)V(1;N)-V(t;N)\right] \\
= \; & \gamma\left[\sum_{x=1}^{t-1}\beta^x\left(F(x,N)-F(x,\gamma N)\right)\right]+ \\
& +(1-\gamma)\left[\sum_{x=1}^{t-2}\beta^x\left(F\left(i_p+1+x,N\right)-F\left(i_p+1+x,(1-\gamma)N\right)\right)\right]+ \\
& +\beta^{t-1}\left[(1-\gamma)\left(F(0,N)-F(t-1,\gamma N)\right)+\gamma\left(F\left(t-1,N\right)-F(t-1,\gamma N)\right)\right] \\
& +\beta^t(1-\gamma)\left[V(1;N)-V(t;N)\right] \\
= \; & \gamma\left[\sum_{x=1}^{t-1}\beta^x\left(F(x,N)-F(x,\gamma N)\right)\right]+ \\
& +(1-\gamma)\left[\sum_{x=1}^{t-2}\beta^x\left(F\left(i_p+1+x,N\right)-F\left(i_p+1+x,(1-\gamma)N\right)\right)\right]+ \\
& +\beta^{t-1}\left[(1-\gamma)F(0,N)+\gamma F(t-1,N)-F(t-1,\gamma N)\right]+ \\
& +\beta^t(1-\gamma)\left[V(1;N)-V(t;N)\right]
\end{aligned}$$

where the first, second and fourth terms are always non-negative (equal to zero for $\gamma = 1$ since

that represents the optimal policy in the constrained case) and we can express the second term as a function $h(\gamma) = (1-\gamma)F(0,N) + \gamma F(t-1,N) - F(t-1,\gamma N)$. It is easy to see that $h(0) = F(0,N) - F(t-1,0) = 1 - F(t-1,0) > 0$ by assumption (8), and that $h(1) = 0$. Also $h'(\gamma) = -F(0,N) + F(t-1,N) - NF_2(t-1,\gamma N)$ and $h''(\gamma) = -N^2F_{22}(t-1,\gamma N)$. Thus, a sufficient (but not necessary) condition for $h(\gamma)$ to be positive in the interval $[0,1]$ is to assure concavity, which is equivalent to the condition $F_{22} \geq 0$ as the lemma states[34].

We have shown that the convex combination of two synchronized planners following the same paths as a split group result in a higher payoff. But because $i_p(N)$ is the optimal threshold for a synchronized planner, then this convex combination must yield a payoff that is no higher than the payoff of a planner using that optimal adoption threshold $i_p(N)$. Thus the lemma follows for this case.

Case II: $t = i_p(N)$

In the second case, suppose that the second group just skips one of the adoption cycles altogether, rejoining the group that adopted a new vintage when it is its second turn to adopt. The payoff for this kind of split is given this time by

$$\gamma\left[-c+1+\sum_{x=1}^{i_p-2}\beta^x F(x,\gamma N)\right]+$$
$$+(1-\gamma)\left[F(i_p+1,N)+\sum_{x=1}^{i_p-2}\beta^x F\left(i_p+1+x,(1-\gamma)N\right)\right]+\beta^{i_p-1}\left[-c+V(0;N)\right]$$

and the convex combination of two planners following similar paths gives a payoff equal to

$$\gamma\left[-c+1+\sum_{x=1}^{i_p-2}\beta^x F(x,N)\right]+$$
$$+(1-\gamma)\left[F(i_p+1,N)+\sum_{x=1}^{i_p-2}\beta^x F\left(i_p+1+x,N\right)\right]+\beta^{i_p-1}\left[-c+V(0;N)\right]$$

the difference between the two given by

$$\gamma\left[\sum_{x=1}^{i_p-2}\beta^x\left(F(x,N)-F(x,\gamma N)\right)\right]+(1-\gamma)\left[\sum_{x=1}^{i_p-2}\beta^x\left(F\left(i_p+1+x,N\right)-F\left(i_p+1+x,(1-\gamma)N\right)\right)\right]$$

which is positive because of the assumption about positive externalities in the function F.

[34] Note in particular that we cannot assure that $h'(\gamma) < 0$ for all values of γ and t, since for t low enough, $F(t-1,N) > F(0,N)$.

Case III: $t = ki_p(N) + l$, $k, l \in \mathcal{Z}$, $l \leq i_p(N)$

This last case represents a generalization of the previous two. In particular, the first case is given by $k = 0$ and case II is given by $k = 1$, $l = 0$. The same logic applied to the previous two cases goes through for this more general case. In particular, if $l \neq 0$ and $k \geq 1$ (which means the group that is not innovating rejoins the other group after more than one cycle) that only represents delaying the adoption cost in case II for l more periods.

In case I we saw that delaying adoption for t periods was not profitable. In this case we are delaying adoption for an extra l periods over case II, which gives the same benefits as in case I, but the losses are bigger: the planner now keeps an older machine as compared with case I (older by $i_p(N)$ vintages), so of course the benefits are lower. This argument in effect indicates that, for given $l \neq 0$ and k, if a split is not profitable, then it is not profitable when using l, $k+1$.

The same argument applies for the case where $l \neq 0$: if it is not profitable to skip k cycles, then it is not profitable to skip $k+1$ cycles either: the cost savings are the same as in case II, but the losses in forgone payoffs are bigger now since the group that delays adoption starts the extra cycle of separation with a machine that is $ki_p(N)$ vintages older that in case II.

This inductive argument shows that no other split is more profitable than the (already proven suboptimal) deviations in cases I and II.

Of course, since the argument in all these three cases does not depend on the size of the group analyzed, no split into more that two groups is profitable, since it can always be decomposed into multiple binary splits. ■

A.2 Proof of Lemma 1

In order to prove that these conditions define an equilibrium threshold i we need to prove that there is no profitable deviation by a single player from this common adoption policy. We will consider first the case of a deviator that decides to delay on the adoption of a different vintage, and then the case of an agent who decides to switch to a different vintage before the time prescribed by the synchronized policy.

Before entering into the details of those deviations, it would be useful to know the valuation of an agent who happens to be outside the synchronized group: suppose all the agents but one are following a synchronized policy of innovating to a new machine every i periods (or equivalently, when their machine reaches an age i). Denote by $V(x, y)$ the value for that last agent of having a

machine of age x when the group has a machine of age y. then we can show that:

Lemma 2 $V(x_0, y) \leq V(x_1, y)$ *for* $x_1 < x_0$, $x_0 \neq y$

Proof. Call agent A the one having a machine of age x_1 and agent B the owner of a machine of vintage x_0. For $x_1 < x_0 \neq y$, $F(x_1, \cdot) > F(x_0, 0)$ because of a lower age (and maybe because of the externalities if $x_1 = y$). Afterwards, agent A can always follow B's policy and get each period a payoff that is not lower than that of B[35]. Note that the condition $x_0 \neq y$ is important since it could be the case that $F(x_0, N) > F(x_1, 0)$ because of the network externalities. ■

Lemma 3 $V(y, y) < V(x, y)$ *for* $0 < x < y - \hat{t}_N$

Proof. Agent A has now a machine of age $x < y - \hat{t}_N$ and agent B owns a machine of age y, the same the group has. Since $y - x > \hat{t}_N$ then $F(y - x, N) < 1 = F(0, 0)$ because of the definition of $\hat{t}_N$ in (10). Moreover, because of assumption (6) about the relative profitability of two machines, we have $F(y, N) < F(x, 0)$. From that point on, agent A can follow the same innovation policy B uses and get each period a payoff that is not lower. ■

With the use of these two lemmas we can show now that no deviation from a synchronized adoption threshold is profitable:

Deviation I: delay for $t \leq \hat{t}_N$ periods

Without loss of generality, set time equal to 0 at the moment where all agents but one adopted a new machine . It is obvious that up to period $\hat{t}_N$ any deviator that decided to keep his machine in period 0 only has two meaningful options: keep his machine one more period or adopt the same vintage the rest of agents in the group are using. The third option, adopting a new machine is clearly worse than adopting the same machine the group is using: since $N > n_l$ and the age of the machine the group is using is no bigger than $\hat{t}_N$, then $F(x, N) > 1 = F(0, 0)$ which is what a single agent can obtain by taking a new vintage in this stage. Moreover, the deviator that joins the group could follow the same policy of the deviator that takes a new vintage and get a payoff that is no lower than the latter.

Having argued that a deviator who delays at $t = 0$ has only two meaningful choices between 0 and $\hat{t}_N$ (either delay until $\hat{t}_N$ or delay and rejoin the group), the first condition establishes that the

[35] See assumption (6)

cost savings from delaying adoption for *one* period and then switching to the vintage used by the rest of agents (the right side of (19)) is smaller than the forgone payoff externality, given by the difference in the left side of (19).

If this condition (19) is satisfied, it will not be profitable to delay at time 0 and "rejoin" the group before time $\hat{t}_N$ either, since the (per-period) loss in payoff for any extra period of delay is greater than $(1-\beta)c$[36].

Deviation II: early adoption: Suppose now that at time $t = 0$ all agents innovate to a new machine and all agents but one will wait until $t = i$ to innovate again. Consider an agent who innovates before age i. Clearly that agent would not innovate —and thus separate from the group—before the age of his machine is $\hat{t}_N$ since it can always wait one more period to switch to that new vintage, saving $(1-\beta)c$ in adoption costs and receiving a bigger payoff during that period by remaining with the group. Left to show is that that agent would not want to adopt a new machine after $\hat{t}_N$ and before age i.

Suppose this agent decides to innovate at time $t \in \{\hat{t}_N + 1, \ldots, i-1\}$. It is obvious that this agent would not want to rejoin the group before time $t = i$ because of lemma 3. If he decides to rejoin the group before $i + \hat{t}_N$ then the best time to do it is at time $\tau(t)$ specified in the proposition. Condition (21) then specifies that the gains of this type of deviation do not compensate the extra adoption costs and the lost payoff (if any).

Since this condition is specified for the maximum over the values of t between 0 and i, then the deviator would never wait more than one complete i cycle to return to the group or to continue alone in cycles of periodicity i. In the first case we have seen that condition (21) makes that deviation unprofitable. In the second case, going alone in cycles of the same periodicity of the group is also unprofitable since the deviator pays the adoption cost before the group does so and does not get the gains from the positive externality of using the group's vintage.

Deviation III: delay for $t > \hat{t}_N$ periods

If we have an agent delaying adoption for more than $\hat{t}_N$ and less than i periods, then because of lemma 2 it would never innovate to the same vintage the group is using when its age is beyond $\hat{t}_N$, since the value of that machine is lower than the value of a new machine. Then left to show is than an agent has no incentive to innovate to a new machine either after $\hat{t}_N$.

[36] Note that the deviator that delays at time 0 receives a payoff below $F(i+1, N)$, whereas the agents that coordinate on adopting a new machine at time 0 get a payoff above $F(0, N)$ until time $\hat{t}_N$. Thus, if (19) is satisfied, then the difference between the two payoffs is greater than $(1-\beta)c$, at least until period $\hat{t}_N$.

Suppose an agent delays adoption until $t \in \{\hat{t}_N + 1, \ldots, i\}$ and adopt a new machine afterwards. Condition (20) says that until time $\hat{t}_N$ this deviation is unprofitable, and afterwards, this deviator would get a payoff lower than that of a deviator of type II that starts at $\hat{t}_N$, so the combined payoff of this deviation is lower than staying with the group.

Left to show is the case where a deviator delays adoption for more than i periods. In that case, we can always decompose that deviation into two parts: the first part is a number ki of periods ($k \in \mathcal{N}$) where the agent just keeps his machine; followed by a second part, where this agent delays for $t < i$ more periods and then adopts a different vintage (as we saw before the only relevant choices could be a new machine or the same machine the rest of agents is using).

The payoff of the first part of the deviation is below that of staying with the group: because of lemma 3 there is no gain for the first i periods and because this deviator gets a lower payoff than in Deviation I for the next cycles (he starts with a lower payoff when considering delay in successive cycles). For the last t periods the payoff is also lower than the corresponding payoff of the agents in the group by an argument similar to that in Deviation I and for the case $t \in \{\hat{t}_N + 1, \ldots, i\}$

These three types of deviations exhaust all the possibilities open to a single agent facing a synchronized group, and thus, no single agent find it optimal to deviate with respect to the group's synchronized policy. ■

A.3 Proof of Proposition 8

In order to prove this proposition, consider a synchronized equilibrium in which agents use the same threshold age $i < \infty$ for adopting a new machine. This staggered strategies generate a steady state distribution that replicates itself every i periods. Denote the period t distribution by $n_t = \{n_{0t}, n_{1t}, n_{2t}, \ldots, n_{it}\}$, where $n_{xt} \geq 0$ and $\sum_{x=0}^{i} n_{xt} = N$.

Suppose at time 0 we introduce an aggregate shock that shifts the entire distribution n_0 by l_θ "steps", meaning that every vintage increases its age by l_θ units. Suppose also that vintages j and k are the two oldest in use by a positive measure of agents at the time the shock hits, that is

$$\begin{aligned} j &= \max\{x : n_{x0} > 0\} \\ k &= \max\{x \neq j : n_{x0} > 0\} \end{aligned}$$

At time 0, the subgroup with the oldest technology j still has to wait for $i - j$ periods to update to the state of the art technology, whereas, the next to thew oldest vintage still needs $i - k$ periods to be updated. Suppose now that, after the shock of size l_θ the two subgroups who were using vintages k and j do not adopt a new machine. We have three possible cases:

Case I: Subgroup j adopts, Subgroup k does not adopt.

In this case we have in period 0 a subgroup of size n_j with vintage 0 and a subgroup of size n_k with vintage $k + l_\theta$. Suppose that group k will adopt a new machine t periods later. An agent would like to switch from group k to group j if a condition similar to (21) holds:

$$\sum_{x=0}^{t} \beta^x \left[F(x, n_j) - F(k + l_\theta + x, n_k)\right] - c(1 - \beta^{t+1}) > 0 \quad (23)$$

In words, inequality (23) assures that an agent will have an incentive to switch to subgroup j when they adopt a new vintage, stay with that subgroup, and then return to subgroup k when they adopt a new vintage.

Since we have assumed that cycles are of finite length, this implies that agents have to adopt at some point, which means that we require $F(0, n) > c(1 - \beta)$ to prevent delay always to occur.

On the other hand, since all groups were using i as an equilibrium threshold before the shock, we can conclude that the no delay condition was satisfied, for all subgroups, namely,

$$1 - F(i + 1, n_l) > (1 - \beta)c$$

where $n_l = \min\{n_{0t}, n_{1t}, n_{2t}, \ldots, n_{it}\}$ is the size of the smallest subgroup of the distribution of agents over vintages. This implies that

$$\begin{aligned} c(1 - \beta^{t+1}) &< [1 - F(i + 1, n_l)] \sum_{x=0}^{\infty} \beta^x \\ &< \sum_{x=0}^{t} \beta^x \left[F(x, n_j)\right] \\ &= \lim_{l_\theta \to \infty} \sum_{x=0}^{t} \beta^x \left[F(x, n_j) - F(k + l_\theta + x, n_k)\right] \end{aligned}$$

Which means that there is a number l_θ big enough to satisfy inequality (23).

Case II: Subgroup j does not adopt, Subgroup k adopts.

In this case the condition for an agent to be willing to switch from group j to group k following

the same deviation as in the previous case is

$$\sum_{x=0}^{t} \beta^x \left[F(x, n_k) - F(j + l_\theta + x, n_j)\right] - c(1 - \beta^{t+1}) > 0$$

and we can use the same argument as before to assure that this inequality is satisfied for values of l_θ big enough.

Case III: No subgroup adopts.

In this case, take group j and suppose that an agent in this group adopts a new machine alone and keeps it until subgroup j adopts a new vintage t periods later and then rejoins them at that point. The net payoff of this deviation is given by

$$\sum_{x=0}^{t} \beta^x \left[F(x, 0) - F(k + l_\theta + x, n_j)\right] - c(1 - \beta^{t+1}) > 0$$

which again is positive for high values of l_θ. A similar argument goes through for subgroup k. ■

References

[1] Adsera, Alicia and Debraj Ray (1997): "History and Coordination Failure," mimeo, Boston University.

[2] Beckmann, Martin (1968): *Dynamic Programming of Economic Decisions*, Springer-Verlag.

[3] Bertola, G. and Ricardo Caballero (1990): "Kinked Adjustment Costs and Aggregate Dynamics," NBER Macroeconomics Annual, 237-287.

[4] Caplin, Andrew and Daniel Spulber (1987): "Menu Costs and the Neutrality of Money", *Quarterly Journal of Economics* 102(4), 703-726

[5] Cooper, Russell and John C. Haltiwanger, Jr. (1992): "Macroeconomic Implications of Production Bunching: Factor Demand Linkages," *Journal of Monetary Economics* 30(1), 107-127.

[6] Cooper, Russell, John C. Haltiwanger, Jr. and Laura Power (1995): "Machine Replacement and the Business Cycle: Lumps and Bumps," Industry Studies Program Working Paper # 62, Boston University.

[7] Fudenberg, Drew and Jean Tirole (1991): *Game Theory*, MIT Press.

[8] Gale, Douglas (1995): "Dynamic Coordination Games," *Economic Theory* 5(1), 1-18.

[9] Gale, Douglas (1996): "Delay and Cycles," *Review of Economic Studies* 63(2) 169-98.

[10] Jovanovic, Boyan and Yaw Nyarko (1996): "Learning by Doing and the Choice of Technology," *Econometrica* 64(6), 1299-1310.

[11] Krugman, Paul R. (1991): "History versus Expectations," *Quarterly Journal of Economics* 106(2), 651-67.

[12] Matsuyama, Kiminori (1991), "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium," *Quarterly Journal of Economics* 106(2), 617-50.

[13] Ruiz, Juan (1998): *Dynamic Interactions in the Presence of Network Externalities*, Ph.D. Thesis. Boston University.

[14] Shleifer, Andrei (1986): "Implementation Cycles," *Journal of Political Economy* 94(6), 1163-1190.