RESPONDERS VERSUS NONRESPONDERS:

A NEW PERSPECTIVE ON HETEROGENEITY

by

John Haltiwanger
Department of Political Economy
The Johns Hopkins University
Baltimore, MD 21218

and

Michael Waldman
Department of Economics
University of California, Los Angeles
Los Angeles, CA 90024

UCLA Working Paper 436
February 1987

ABSTRACT

This paper considers the implications of a particular type of heterogeneity - one which characterizes a large number of economic environments, but which has not received any systematic treatment in the literature. We refer to this heterogeneity as responders versus nonresponders. In this paper we provide a general method of analysis for this heterogeneity, and then show how this general method of analysis can be used to understand a wide variety of economic environments. Particular applications considered include: (i) the recent work on the evolution of market outcomes given network externalities; (ii) recent research on heterogeneity in information processing abilities; and (iii) work on reputation in models exhibiting the last period problem.

I. Introduction

In this paper we consider the implications of a particular type of heterogeneity — one which characterizes a large number of economic environments, but which has not received any systematic treatment in the literature. We refer to this heterogeneity as responders versus nonresponders. The term responders here refers to agents whose behavior depends on actions chosen by other agents in the population, while nonresponders are agents whose behavior is independent of the actions chosen by others. In considering environments characterized by heterogeneity of this sort, our focus will be on the question — is it the responders or is it the nonresponders who are disproportionately important? That is, relative to the numbers of the two types of agents in the population, does the equilibrium more closely resemble what occurs when everyone is a responder or does it more closely resemble what occurs when all are nonresponders? We feel the answer to this question provides valuable insights into the nature of equilibrium across a wide range of economic settings.

The significance of our approach is evidenced by the number of recent important and quite diverse developments in the literature which can be understood in terms of responders versus nonresponders. One such development is the recent work on the evolution of market outcomes in the presence of "network externalities" (see e.g., David (1985), Farrell and Saloner (1985, 1986), and Katz and Shapiro (1986)). One of the main results in this literature is that, when network externalities are present, there tends to be a first mover advantage. That is, if one technology is superior in the early stages of an industry's evolution, there is a tendency for that technology to remain the standard even after it is no longer superior. What we will demonstrate is that this result follows from the perspective of responders versus nonresponders. The logic is that consumers who purchase prior to the introduction of a new technology are nonresponders. That is, their behavior is independent of the actions of consumers who purchase after the new technology is made available. In addition, network externalities, or what we will refer to as synergism, is exactly the property which causes nonresponders to be disproportionately important. The result is that agents who purchase prior to the introduction of the new technology have a disproportionate impact on equilibrium, which in turn explains why the initial technology may remain the standard even after a superior new technology has become available.

A second development in the literature which can be viewed in terms of responders versus nonresponders is the recent research on heterogeneity in information processing abilities (see e.g., Conlisk (1980), Akerlof and Yellen (1985a,b), Haltiwanger and Waldman (1985, 1986a), and Russell and Thaler (1985)). For example, in our own earlier work we considered whether rational expectations equilibria are robust to the introduction of agents who do not satisfy a rational expectations assumption. The way we approached the problem was to ask the following question. In a world which contains both agents who satisfy a rational expectations assumption and agents who are more limited in their ability to form expectations, is it the rational expectations agents or is it the more naive agents who are disproportionately important in the resulting equilibrium? What we would now argue is that many of the results we derived in this earlier context can be understood in terms of responders versus nonresponders. The logic is that agents with rational expectations typically fit the description of what we are here referring to as responders, while in many cases the more naive agents fit the description of what we are here referring to as nonresponders.

The third recent development which can be viewed in terms of responders versus nonresponders is the pioneering work on reputation of Kreps and Wilson (1982), and Milgrom and Roberts (1982). The issue they are concerned with is the last period problem. That is, in a finite period game theoretic model, it is difficult to find a role for reputation building because the incentive to "cheat" in the last period causes any equilibrium consistent with reputation building to unravel.[1] They demonstrate that by adding a little bit of uncertainty concerning agents' preferences, the last period problem can be avoided and equilibria consistent with reputation building are then easy to construct. What we will argue is that this result can also be understood in terms of responders versus nonresponders. That is, when these authors introduce uncertainty, from our standpoint they are simply introducing a probability that agents are nonresponders whose behavior is to not cheat. Further, the environment considered in these papers is one in which nonresponders are disproportionately important, with the result being that reputation building, i.e., not cheating, is standard behavior.

The goal of the present paper is threefold. First, we would like to present a general method of analysis for the heterogeneity we refer to as responders versus nonresponders. In particular, we would like to develop a framework which will allow us to consider any environment characterized by this heterogeneity, and quickly answer whether in that environment it is the responders or the nonresponders who are disproportionately important. Second, we would like to demonstrate that this heterogeneity can arise both in a large number of settings, and for a variety of quite diverse reasons. Examples of the latter include differences in tastes across agents, differences in the ability of agents to process information, and differences across agents in the dates in which actions are undertaken. Third, we would like to show how our approach can be used to understand the three recent developments in the literature discussed above – as well as use the approach to provide new results concerning both these developments and other important issues in the literature.

The plan of the paper is as follows. In Section II we present our general method of analysis for the heterogeneity we refer to as responders versus nonresponders. In Sections III-V we then consider three different environments which are characterized by this heterogeneity, and show how our approach can be used to help analyze these environments. Section III considers an environment where the heterogeneity arises becauses of differences in the dates in which actions are undertaken. The specific model considered is closely related to the models discussed earlier of Farrell/Saloner and Katz/Shapiro. Section IV considers an environment where the heterogeneity arises because of differences in the ability of agents to process information. In particular, we consider a model which is in the spirit of Akerlof and Yellen (1985a,b), and show the relationship between our approach of responders versus nonresponders and their work concerning maximizers and nonmaximizers. Section V considers an environment where the heterogeneity arises because of differences in tastes across agents. Here we consider a labor market setting where workers have the opportunity to shirk, while the firm offers a non-market clearing wage in order to deter shirking. In this section we also discuss how our approach can be helpful in understanding the earlier mentioned work on reputation of Kreps/Wilson and Milgrom/Roberts. Finally, Section VI contains some concluding remarks.

II. The General Approach

In this section we present a general approach to the analysis of environments characterized by what we refer to as responders versus nonresponders. As stated earlier, the term responders refers to agents whose behavior is dependent on the actions chosen by other agents in the population, while the term nonresponders refers to agents whose behavior is independent of the actions chosen by others. Also, recall that our general approach to the problem is an attempt to answer the following question. Given an environment in which both responders and nonresponders are present, is it the responders or is it the nonresponders who are disproportionately important?

The crucial factor in answering the above question is the nature of the interaction among agents. Consider first a world which displays what we will refer to as congestion. That is, for any agent i, the larger is the total number of agents who choose a particular behavior, the lower is the incentive for agent i to choose that behavior. For this type of environment it is the responders who are disproportionately important. In other words, the equilibrium will more closely resemble what occurs when all agents are responders than would be suggested by the relative number of responders and nonresponders in the population.

The logic behind this result is as follows. Suppose that all agents choose between two actions – denoted actions A and B. Further, to keep it simple suppose that if all agents were responders there would be some agents choosing each action, while if all were nonresponders then everyone would choose action A. We can now think about what occurs when the population consists of a mix of the two types. By definition the actions of nonresponders in the mixed case is exactly the same as their actions in the pure nonresponder case, i.e., all continue to choose action A. Now compare the actions of responders in the mixed case with their actions in the pure responder case. Because of congestion, the presence of nonresponders lowers the incentive for agents to choose action A, and raises the incentive for action B. The result is that responders alter their behavior by having a larger proportion choose action B. In other words, the responders are disproportionately important given congestion, because the presence of nonresponders is somewhat nullified by the "response" of the responders to their presence.

Now consider a world which displays what we will refer to as synergism. That is, the larger is the total number of agents who choose a particular behavior, the <u>higher</u> is the incentive for agent i to choose that behavior.[2] For this type of environment it is the nonresponders who are disproportionately important. In other words, the equilibrium will now more closely resemble what occurs when all agents are <u>nonresponders</u> than would be suggested by the relative number of the two types in the population.

The logic behind this result is related to that given above. Suppose again that all agents choose between actions A and B, and that if all agents were responders there would be some agents choosing each action, while if all were nonresponders then everyone would choose action A.[3] We can again compare the actions of responders in the mixed case with their actions in the pure responder case. Because of synergism, the presence of nonresponders now raises the incentive for agents to choose action A, and lowers the incentive for action B. The result is that responders alter their behavior by having a larger proportion choose action A. In other words, the nonresponders are disproportionately important given synergism, because the responders now "respond" in a manner which reinforces, rather than nullifies, the behavior of the nonresponders.

In summary, whether it is the responders or the nonresponders who are disproportionately important depends on the nature of the interaction among agents. If the incentive to choosing a behavior is negatively related to the total number of agents who choose that behavior, i.e., if the environment exhibits what we refer to as congestion, then it is the responders who are disproportionately important. The reason is that, in this case, the presence of nonresponders is somewhat nullified by the manner in which responders alter their behavior. On the other hand, if the incentive to choosing a behavior is positively related to the total number of agents who choose that behavior, i.e., if the environment exhibits what we refer to as synergism, then it is the nonresponders who are disproportionately important. The logic here is that the presence of nonresponders now causes responders to alter their behavior in a manner which reinforces, rather than nullifies, the presence of the nonresponders.

III. Application 1: Evolution of Market Outcomes Given Network Externalities

In this section we consider an environment in which responders and nonresponders are present because of differences across agents in the dates in which actions are undertaken. In particular, we consider a model which is typical of the literature mentioned in the introduction on the evolution of market outcomes given network externalities. Remember, the concept of network externalities is basically the same as what we refer to as synergism (see footnote 2). Two main results have come out of this body of literature. David (1985) and Katz and Shapiro (1986) have stressed the idea of a first mover advantage, i.e., a technology which is superior in the early stages of an industry's evolution tends to remain the standard even after it is no longer superior. On the other hand, Farrell and Saloner (1986) demonstrate that from a social welfare standpoint there may be either excess inertia (inefficient nonadoption of a new technology), or excess momentum (inefficient adoption of a new technology). Although at first glance the two results may seem somewhat inconsistent, they are not. In this section we will show how the first mover advantage result follows from the perspective of responders versus nonresponders, and then use the intuition we develop to discuss the results of Farrell and Saloner. The particular model we analyze is most closely related to one considered by Farrell and Saloner.[4]

Our model is an overlapping generations model in a continuous time setting. Agents are infinitesimal and arrive over time at a constant flow rate n(t), where each agent lives for exactly L periods. Let $\bar{N}=\int_0^L n(t)dt$, i.e., $\bar{N}$ is the total number of agents alive at any moment of time. Except for the dates in which they are born, all agents are identical. The problem faced by an agent is the choice of which product to consume. Prior to date T^* only product X is available, where X is produced by a perfectly competitive industry at constant marginal cost c_x, i.e., the price of X is c_x. At date T^* product Y becomes available, where this innovation is not anticipated by agents born prior to date T^*.[5] Product Y is also produced by a perfectly competitive industry, where its constant marginal cost of production is c_y. Finally, products X and Y are substitutes, i.e., at any date in time an individual can derive benefits from one or the other, but not both.

We can now consider the choices available to an agent born in period t'. First, he can decide not to purchase either good, in which case his utility equals 0. Second, he can decide to purchase X at some date $\hat{t}$, $t'\leq\hat{t}<t'+L$, in

which case his utility equals $\int_{\hat{t}}^{t'+L} F_x(N_x^t)dt - c_x$, where $F_x'>0$ and N_x^t denotes the total number of agents alive at date t who have purchased X. Third, if Y is available then he may decide to purchase Y at some date $\hat{t}$, $t' \leq \hat{t} < t'+L$, in which case his utility equals $\int_{\hat{t}}^{t'+L} F_y(N_y^t)dt - c_y$, where $F_y'>0$ and N_y^t denotes the total number of agents alive at date t who have purchased Y.[6] The assumptions $F_x'>0$ and $F_y'>0$ capture that network externalities are present or the idea that the environment exhibits synergism.

In their model, Farrell and Saloner impose a condition which guarantees that agents who purchase X prior to date T^* do not purchase Y when Y becomes available (see their footnote 10). This is an important condition to impose in that, given the condition, their model fits our framework of responders versus nonresponders. That is, agents born prior to date T^* are nonresponders. They purchase X when they are born, and independent of the behavior of agents born T^* or later, they do not purchase Y when Y becomes available. On the other hand, agents born T^* or later are responders. That is to say, their behavior does depend on the actions chosen by other agents in the economy.

To guarantee that our model exhibits the property discussed above we assume $LF_x(0) > LF_y(\bar{N}) - c_y$. Intuitively, this says that if X is available to a consumer at no cost, e.g., if he has already purchased it, then the consumer will not have an incentive to purchase Y.[7,8]

We can now proceed to the analysis. The first step is to establish a benchmark with which later results can be compared. In particular, we will consider what equilibrium looks like when both X and Y are available in all periods. The way we will proceed is to hold everything in the model fixed except for c_x, and then see how equilibrium depends on the value for c_x. Note, following Farrell and Saloner we restrict attention to equilibria where, given a date in which both X and Y are available, for all later dates either all consumers choose X or all consumers choose Y (other equilibria require quite implausible behavior on the part of agents).

<u>Proposition 1</u>: Suppose both X and Y are available in all periods. Then there will be values $\bar{c}$ and $\underline{c}$, $\bar{c}>\underline{c}$, which satisfy the following.

i) If $c_x<\underline{c}$, then the unique equilibrium is that all agents purchase X at birth.

ii) If $c_x>\bar{c}$, then the unique equilibrium is that all agents purchase Y at birth.

iii) If $\underline{c}\leq c_x\leq\bar{c}$, then there are two equilibria. One equilibrium is that all agents purchase X at birth, while the other is that all agents purchase Y at birth.

Proposition 1 is straightforward. First, if the cost of producing X is very low ($<\underline{c}$), then X is quite attractive and the unique equilibrium is that all agents purchase X. Second, if the cost of producing X is very high ($>\bar{c}$), then X is unattractive and the unique equilibrium is that all agents purchase Y. Third, there is an intermediate range of values for the cost of producing X ($\underline{c}\leq c_x\leq\bar{c}$) such that both of the above possibilities are equilibria. This final result is not surprising since it is already well established in the literature that multiple equilibria can exist when synergism is present (see Farrell and Saloner for a more detailed discussion of why multiple equilibria can exist in this setting).

The next step of the analysis is to go back to the original specification where Y only becomes available at date T^*, and again see how equilibrium depends on the value for c_x.[9]

<u>Proposition 2</u>: Suppose X is available in all periods, but Y only becomes available at date T^*. Then there will be values $\bar{c}'$ and $\underline{c}'$, $\bar{c}'>\underline{c}'$, which satisfy the following.

i) If $c_x<\underline{c}'$, then all agents purchase X at birth.

ii) If $c_x>\bar{c}'$, then agents born prior to date T^* purchase X at birth and agents born T^* or later purchase Y at birth.

iii) If $\underline{c}'\leq c_x\leq\bar{c}'$, then each of the above are equilibria.

iv) $\underline{c}'>\underline{c}$ and $\bar{c}'=\bar{c}$.

Proposition 2 tells us that, as in Proposition 1, there are two critical values for c_x. When c_x is below $\underline{c}'$ then all agents continue to purchase X after Y becomes available. When c_x is above $\bar{c}'$ then starting at date T^* all newly born agents purchase Y. When c_x is between these two critical values, then both possibilities are equilibria.

What we find to be the most interesting aspect of Proposition 2 is the comparison between $\underline{c}'$ and $\underline{c}$. The comparison tells us that there is a range of values for c_x with the following properties. First, when both products are available in all periods it is an equilibrium for all agents to purchase Y. Second, when Y only becomes available at date T^* no Y is ever purchased. This result follows from the perspective of responders versus nonresponders. Agents born prior to date T^* are nonresponders and this is an environment which exhibits synergism. Section II, therefore, tells us that the agents born prior to date T^* should be the ones who are disproportionately important. The result is that there is a range of values for c_x where everyone purchasing Y is an equilibrium when it is available in all periods, but Y is <u>never</u> purchased when it only becomes available at date T^*.[10]

Although Proposition 2 is of interest, it does not directly get at the result mentioned earlier concerning a first mover advantage, i.e., a technology which is superior early in an industry's evolution tends to remain the standard even after it is no longer superior. The reason is the range of values for c_x for which there are multiple equilibria. In what follows we present a particular way of resolving this multiple equilibria problem, and then consider the idea of a first mover advantage in the context of what results.

Think back to the situation where both X and Y are available in all periods. For that case we spoke of two possible equilibria: one where all agents purchase X and one where all agents purchase Y. Given this environment it is easy to find a critical value $\hat{c}$ which has the following properties. First, if $c_x<\hat{c}$, then every agent would prefer an equilibrium where all agents purchase X over one where all agents purchase Y. Second, if $c_x>\hat{c}$, then every agent would prefer an equilibrium where all agents purchase Y over one where all agents purchase X (note: if $c_x=\hat{c}$, then all agents would be indifferent between the two equilibria).

Now consider again the situation where Y only becomes available at date T^*. One might suggest that when multiple equilibria are present, the one which is most likely to result is the one which is pareto preferred. Unfortunately, even if we restrict the pareto criterion to agents born T^* or later, this suggestion does not resolve the problem. For some values of c_x agents born near date T^* will prefer the equilibrium where only X is purchased, while agents born later will prefer the equilibrium where the economy switches over to purchasing Y.

Our alternative suggestion is that, when multiple equilibria are present, the equilibrium most likely to result is the one which is preferred by agents born at date T^*. The argument follows. Suppose that these agents prefer the equilibrium where all agents purchase X (a similar argument holds for the other case). If these agents purchase X, then there will be multiple equilibria for the sub-game starting at date $T^*+\epsilon$. However, agents born at date $T^*+\epsilon$ would prefer the sub-game equilibrium where they purchase X over the one where they purchase Y. That is, the agents born at date $T^*+\epsilon$ would have no incentive to move the economy away from the equilibrium begun at date T^*. Further, this argument can be repeated over and over to show that no agents born after date T^* would have an incentive to move the economy away from this equilibrium.

Now suppose that agents born at date T^* purchase Y, even though they prefer the equilibrium where X is purchased. There are again multiple equilibria for the sub-game starting at date $T^*+\epsilon$. In addition, the agents born at date $T^*+\epsilon$ again prefer the sub-game equilibrium where they purchase X over the one where they purchase Y. That is, the agents born at date $T^*+\epsilon$ now do have an incentive to move the economy away from the equilibrium begun at date T^*.

Overall, this suggests to us that the most plausible equilibrium is the one preferred by agents born at date T^*. The logic from the above discussion is that if these agents choose the equilibrium they prefer, then no future set of agents will have an incentive to move the economy away from this equilibrium. On the other hand, an analogous property is <u>not</u> satisfied if the agents born at date T^* choose the equilibrium they do not prefer.

In the following proposition we analyze what happens when Y only becomes available at date T^*, and multiple equilibria are resolved in the above suggested manner.

<u>Proposition 3</u>: Suppose X is available in all periods, but Y only becomes available at date T^*. Further, suppose that when multiple equilibria are present, the equilibrium which results is the one preferred by agents born at date T^*. Then there will be a value $\hat{c}'$ which satisfies the following.

i) If $c_x<\hat{c}'$, then all agents purchase X at birth.

ii) If $c_x>\hat{c}'$, then agents born prior to date T^* purchase X at birth and agents born T^* or later purchase Y at birth.

iii) $\hat{c}'>\hat{c}$.

Not surprisingly, Proposition 3 tells us that if c_x is below a critical value then only X is purchased, while if it is above, then starting at date T^* only Y is purchased. More interestingly, the proposition exhibits the first mover advantage discussed earlier. Recall that $\hat{c}$ is defined as follows. If both products are available in all periods and $c_x<(>)\hat{c}$, then every agent would prefer an equilibrium where all agents purchase X(Y) over one where all agents purchase Y(X). Given this, iii) tells us that there is a first mover advantage. That is, there is a range of values for c_x such that Y is the superior product (as defined by $\hat{c}$), yet the equilibrium is that Y is not purchased even after it becomes available. As with the earlier comparison between $\underline{c}'$ and $\underline{c}$, this result follows from the perspective of responders versus nonresponders. Because this environment exhibits synergism and because the agents born prior to date T^* are nonresponders, it is these agents who should be disproportionately important. This is manifested in Proposition 3 in that, when Y becomes available, X may remain the standard even when it is not the superior product.

We can now discuss the results of Farrell and Saloner. As indicated earlier, their main finding was that from a social welfare standpoint there may be either excess inertia (inefficient nonadoption of a new technology), or excess momentum (inefficient adoption of a new technology). Although at first glance this may seem inconsistent with our demonstration of a first mover advantage, it is not. The two results are compatible because, from a societal standpoint, some type of first mover advantage is beneficial. The reason is that there are transition costs associated with an economy switching from X to Y — the costs being due to the loss in welfare incurred by consumers who purchase X prior to Y becoming available (Farrell and Saloner refer to this as stranding). Further, what Farrell and Saloner demonstrate is that relative to the first mover advantage which would be best for society, the actual first mover advantage may be either too large (excess inertia), or too small (excess momentum).

IV. Application 2: Production Externalities and Near Rationality

A critical issue in the modeling of economic behavior is how well can agents process information. The standard approach is to assume that agents are "rational," or equivalently, that agents have unlimited abilities to process information. This has led to a recurring controversy, however, because as has been pointed out by many previous authors, real world agents are obviously limited in these abilities. As a result of this controversy, several alternatives to the rationality assumption have been suggested over time. A prominent example is the concept of satisficing developed by Simon and his followers.[11] More recently, however, work in this area has shifted in a new direction. A number of authors have investigated the idea that agents tend to be heterogeneous in terms of information processing abilities (see the references in the introduction). In particular, these authors have looked at models where one group of agents processes information in a very sophisticated manner, while others are much more limited in their capabilities. In this section we investigate some of the links between this literature and our general approach of responders versus nonresponders.

As already indicated in the introduction, there is a very direct link between our own earlier work and responders versus nonresponders. For example, in Haltiwanger and Waldman (1985) we considered a static environment in which there were both agents who satisfy a rational expectations assumption, and agents who are more limited in their ability to form expectations. We showed that rational expectations equilibria tend to be robust when the environment exhibits congestion, but are not robust when synergism is present. The relationship here is straightforward. Agents with rational expectations tend to be responders, while in a static setting naive agents tend to be nonresponders. Consistent with Section II, therefore, rational expectations equilibria should be less robust in an environment which exhibits synergism.[12]

Another issue of interest is the relationship between responders versus nonresponders and the concept of near rationality introduced by Akerlof and Yellen (1985a,b). Akerlof and Yellen explore the ramifications of having a population consist in part of nonmaximizers, where nonmaximizers are defined as agents who do not respond at all to small shocks. They demonstrate that in many environments the presence of nonmaximizers has a first order impact on the economy, even though the private loss associated with being a nonmaximizer is second order. In situations in which the private loss satisfies this criterion, they denote nonmaximizers as being near rational.

It should be clear that Akerlof and Yellen's nonmaximizers satisfy the definition of what we refer to as nonresponders, i.e., the behavior of a nonmaximizer is independent of the behavior of other agents in the population. Given this, consider an environment where near rational agents make mistakes which are second order in terms of the cost to the individual, but which have a first order effect on the equilibrium. One might conjecture that this first order effect should be larger the more synergistic is the environment. In other words, the first order effect should be larger when nonresponders are disproportionately important.

To investigate the above conjecture we consider a simple model of production where congestion and synergism are potentially present because the model exhibits production externalities (for a discussion of production externalities see Layard and Walters (1978), pp. 224-226). In particular, we assume a continuum of firms identical in terms of costs of production, where N denotes the total number of firms in the continuum. Let x_i denote the output of firm i and X denote aggregate output. Each firm i is assumed to face the cost function given in equation (1).

(1) $$C_i(x_i)=\alpha c(x_i)f(X),$$

where $c(0)=0$, $c'(0)=0$, $c''>0$, and $c'(\infty)=\infty$. Production externalities are captured by the term $f(X)$, where there are negative production externalities if $f'>0$ and positive externalities if $f'<0$. Notice that the negative externalities case is the same as saying the model exhibits congestion, while positive externalities translates into synergism. The logic here is straightforward. When negative externalities are present the more which is produced on average, the smaller is the incentive for any particular firm to have a high output, i.e., congestion. On the other hand, positive externalities means that the more which is produced on average, the larger is the incentive for any particular firm to have a high output, i.e., synergism.[13]

Following in the spirit of Akerlof and Yellen, we assume there are two types of firms. A proportion p of the firms are maximizers, where maximizers have low costs of processing information and thus follow standard assumptions concerning how firms behave. On the other hand, a proportion (1-p) are nonmaximizers, where nonmaximizers have high costs of processing information with the subsequent result being they do not respond at all to small shocks. Finally, it is assumed that the industry faces a perfectly elastic demand curve for its output, where the price of a unit of output is normalized to one.[14]

This completes the set-up of the model, and we can now proceed to the analysis. Suppose that the economy is initially in a long run equilibrium where all agents are exactly maximizing. Let α^* denote the value for α in this long run equilibrium. Propositions 4, 5 and 6 consider what happens if α now rises to $\alpha^*+\epsilon$, where, because it is a small shock, for nonmaximizers x_i remains unchanged. Note, below $\pi^m(\pi^n)$ denotes the profit of a maximizer (nonmaximizer), W denotes social welfare, $W^m(W^n)$ denotes social welfare for the special case where all firms are maximizers (nonmaximizers), and $X^m(X^n)$ denotes aggregate output for the special case where all firms are maximizers (nonmaximizers).[15]

Proposition 4:

(2) $$\left.\frac{d\pi^m}{d\epsilon}\right|_{\epsilon=0} = \left.\frac{d\pi^n}{d\epsilon}\right|_{\epsilon=0},$$

while

(3) $$\left.\frac{d(X-X^m)}{d\epsilon}\right|_{\epsilon=0} > 0$$

and

(4) $$\left.\frac{d(W-W^m)}{d\epsilon}\right|_{\epsilon=0} <(>)0 \text{ if } f'>(<)0.$$

Proposition 4 tells us that this model satisfies the Akerlof and Yellen criterion discussed above.[16] On the one hand, (2) tells us that the private loss associated with being a nonmaximizer is second order. On the other hand, (3) and (4) state that in terms of both aggregate output and social welfare the presence of nonmaximizers has a first order impact on the economy.

The next step is to show that this model is consistent with our own results concerning responders versus nonresponders.

Proposition 5: If $f'>(<)0$, then $X<(>)pX^m+(1-p)X^n$. Further, $\left.\frac{dW}{d\epsilon}\right|_{\epsilon=0} >(>) p\left.\frac{dW^m}{d\epsilon}\right|_{\epsilon=0} + (1-p)\left.\frac{dW^n}{d\epsilon}\right|_{\epsilon=0}$, where $\left.\frac{dW^m}{d\epsilon}\right|_{\epsilon=0} >(<) \left.\frac{dW^n}{d\epsilon}\right|_{\epsilon=0}$.

We know $X^m<X^n$. Hence, Proposition 5 tells us that if $f'>0$, then in terms of both aggregate output and social welfare it is the maximizers who are disproportionately important. However, if $f'<0$, then it is the nonmaximizers who dominate.[17] The intuition for these results follows directly from the discussion of Section II. As stated earlier, negative externalities is the same as saying the model exhibits congestion, while positive externalities translates into synergism. Hence, Section II tells us that maximizers (responders) should be disproportionately important when negative externalities are present, while nonmaximizers (nonresponders) should be disproportionately important given positive externalities.

Now that we see this model satisfies both the results concerning near rationality of Akerlof and Yellen and our own results concerning responders versus nonresponders, we can explore the relationship between the two approaches. To consider this issue we define what will be referred to as an increasing synergistic transformation of f(.). In particular, $\hat{f}(.)$ is an increasing synergistic transformation of f(.) if there exists some value $\tilde{X}$ such that $\hat{f}(\tilde{X})=f(\tilde{X})$ and $\hat{f}'(\tilde{X})<f'(\tilde{X})$. In other words, an increasing synergistic transformation of f(.) varies the slope of f(.) around some fixed point.

<u>Proposition 6</u>: Consider an increasing synergistic transformation of f(.) which leaves the initial long run equilibrium unchanged. Such a transformation causes $\left.\frac{d(X-X^m)}{d\epsilon}\right|_{\epsilon=0}$ to increase. Further, if $f'<0$, then $\left.\frac{d(W-W^m)}{d\epsilon}\right|_{\epsilon=0}$ also increases.

Taken together, Propositions 4 and 6 tell us that, at least in terms of aggregate output, there is a very close relationship between the two approaches. That is, they state that the first order effect identified by Akerlof and Yellen will be larger the more synergistic is the environment. In other words, the first order effect is larger when nonmaximizers (nonresponders) are disproportionately important. On the other hand, in terms of social welfare the results are somewhat mixed. If one starts with an initial situation characterized by synergism ($f'<0$), then increasing the degree of synergism does increase the first order effect. However, if one starts with an initial situation characterized by congestion ($f'>0$), then increasing the degree of synergism can actually cause the first order effect to move in either direction.

The intuition for why there are mixed results for social welfare is as follows. The first order effect for social welfare is equal to the aggregate output first order effect multiplied by the magnitude of the externality. If one starts with an initial situation characterized by synergism, then an increase in the degree of synergism increases both components. The subsequent result, therefore, is an increase in the first order effect for social welfare. On the other hand, if one starts with an initial situation characterized by congestion, then an increase in the degree of synergism causes the first component to rise and by definition the second to fall. Hence, that this case results in an ambiguous change of the first order effect simply means that either of the movements identified above can be the dominant one.

We can now reconsider what one learns about the rationality assumption from our own earlier work and that of Akerlof and Yellen. If one focuses mainly on the results concerning aggregate output, then an interpretation of the above analysis is that the basic message of the two approaches is really quite similar. In our earlier work we considered whether rational expectations equilibria are robust to the introduction of agents who do not satisfy a rational expectations assumption. Our conclusion was that equilibria tend to be robust when the environment exhibits congestion, but are not robust when synergism is present. On the other hand, Akerlof and Yellen consider what happens to equilibria when nonmaximizers are introduced, where nonmaximizers are agents who do not respond at all to small shocks. They demonstrate that even though the private loss to being a nonmaximizer is typically second order, the introduction of such agents will frequently have a first order impact on the economy. Further, what we have shown above is that this first order impact tends to be larger the more synergistic is the environment. Or in summary, both approaches suggest that employing the assumption of rationality is problematic for environments which exhibit synergism.

V. Application 3: Shirking and Non-Market Clearing Wages

In this section we consider an environment where the heterogeneity we refer to as responders versus nonresponders arises because of differences in tastes across agents. The particular environment considered is a labor market setting where workers have the opportunity to shirk, while the firm offers a non-market clearing wage in order to deter shirking.[18] At the end of the section we also discuss a second environment in which responders and nonresponders are present because of differences in tastes, i.e., the earlier mentioned work on reputation of Kreps/Wilson and Milgrom/Roberts.[19]

We consider a single period/partial equilibrium setting which is in the spirit of Becker and Stigler (1974), and Calvo and Wellisz (1979). There is a single firm which wants to hire $\tilde{N}$ workers from a pool of self-employed workers, where it will be convenient to assume that this pool of self-employed consists of a continuum of agents. Following Calvo and Wellisz, we assume that if a worker is hired at the firm, then at the beginning of the employment period the worker makes an irreversible decision concerning whether or not to shirk for the entire period. In particular, if worker i is employed at the firm and he shirks he then produces an amount Y, while if he expends effort he produces an amount X, where $X>Y$.

Each worker i has the following utility function which describes the utility he receives if he is employed at the firm.

(5) $$U_i = U(W_i) - \delta_i \gamma_i k,$$

where $U' > 0$, $U'' \leq 0$, $k < U(X) - U(Y)$, and W_i denotes the wage received by worker i. $\delta_i = 0(1)$ if the worker does (does not) shirk, while γ_i is a draw from a random variable which equals 0 with probability (1-p) and 1 with probability p. That is, for each worker i there is a probability (1-p) that the worker receives no disutility from effort, and a probability p that he receives disutility – where k is the disutility he receives. It is also assumed that whether or not a worker receives disutility from effort is private information to that worker.

The firm is risk neutral, while the only input used in the production process is labor. Further, because it is assumed that only aggregate output is observed, a worker's wage cannot be made directly contingent on his own output. What firms do instead is monitor workers. In particular, again following Calvo and Wellisz (1979), an inspection is carried out at the beginning of the employment period. It is assumed that if worker i is caught shirking, he is then fired and becomes self-employed. In this case his utility level is U(H) if $\gamma_i = 0$ (he does not receive disutility from effort), while it is U(H)-k if $\gamma_i = 1$ (he does receive disutility), where $H < Y$.[20] On the other hand, if he is not caught shirking, he is then employed for the entire period and he receives wage W.

If worker i is shirking, then during the inspection he will be caught with probability m_i. That is, workers vary in terms of their ability to avoid being detected, where it is further assumed that each worker's value for m_i is private information to that worker. The distribution of m_i's in the population is independent of the distribution of γ_i's, and, in particular, is described by a frequency distribution function G(.) defined over the interval

[0,1], where $G'(m)>0$ for all $0<m<1$ and $\lim_{m\to 1} G(m)<1$.[21]

We will now proceed to analyze the above model. If a worker receives no disutility from effort, then clearly he will never have an incentive to shirk. That is, the agents who receive no disutility from effort are nonresponders. Now consider a worker who does receive disutility from effort. In this case the decision concerning whether or not to shirk depends on both the wage offered by the firm and the worker's value for m_i, where the higher is the wage the lower is the incentive for the worker to shirk. For example, if we assume risk neutrality, i.e., $U''=0$, then the worker will (will not) shirk as long as $m_i<(>)k/[W-(H-k)]$. As will become clear, this decision problem is such that the behavior of an agent does depend on the other agents in the population. That is, the agents who receive disutility from effort are responders.

Our first proposition deals with how changes in p affect the wage. Note, all proofs are relegated to an Appendix.

<u>Proposition 7</u>: An increase in p causes W to increase.

Proposition 7 states that an increase in the proportion of workers who receive disutility from effort causes an increase in the wage. The result is not particularly surprising. The way the firm deters shirking is by offering a wage which is above the market clearing level. Hence, if there is an increase in the proportion of workers willing to shirk, it is not surprising that the firm will want to offer a higher wage. The result is important, however, in that it tells us how this model relates back to the discussion of Section II. Specifically, it tells us that this is an environment which exhibits congestion. The more workers willing to shirk, the lower is the return to shirking. The logic is that more potential shirkers leads to a higher wage, and the higher wage lowers the return to shirking. In turn, now that we know this is an environment which exhibits congestion the following proposition becomes easy to understand. Note, in the following $N_S(N_{NS})$ will denote the total number of workers employed by the firm who decide to (not to) shirk, while $\hat{N}_S(\hat{N}_{NS})$ denotes this value for the polar case p=1.

<u>Proposition 8</u>: If $0<p<1$, then $N_S>p\hat{N}_S$ and $N_{NS}<p\hat{N}_{NS}$.

Proposition 8 tells us that in this world it is the agents who receive disutility from effort who are disproportionately important. That is, the amount of shirking which takes place more closely resembles what occurs when all agents receive disutility from effort than would be suggested by the relative number of the two types of agents in the population. This conclusion follows directly from the discussion of Section II. That is, this is an environment which exhibits congestion, and it is also an environment in which the agents who receive disutility from effort are the responders. Hence, Section II tells us that it is these agents who should be disproportionately important.

The above result has important implications for the large body of work in the labor economics literature which abstracts away from the desire of workers to shirk. That is, the above result suggests that one should be very cautious in making this abstraction. The reason is that agents who are willing to shirk are in many cases disproportionately important, and hence, it is quite possible that having even just a small number of them can have a significant impact on the nature of equilibrium.[22]

We will end this section with a discussion of a second environment in which responders and nonresponders are present because of differences in tastes across agents. In their pioneering work on reputation, Kreps/Wilson and Milgrom/Roberts (KWMR) attempt to provide a resolution to the last period problem. The last period problem being that, in a finite period game theoretic setting, it is difficult to find a role for reputation building because the incentive to "cheat" in the last period causes any equilibrium consistent with reputation building to unravel. To deal with this problem these authors consider what Selten (1978) refers to as the chain-store game. The game is described as follows. There is a monopolist who operates in N separate markets, and who sequentially faces a different potential entrant in each market. Further, in each market the potential entrant moves first by deciding whether or not to enter the market, where in making this decision the agent knows the actions which were taken in earlier markets. If the potential entrant decides not to enter, then there are no further moves by players in that market. If entry does occur, however, then the monopolist must decide whether to cooperate or to act aggressively. Finally, payoffs satisfy the following three conditions. First, if entry has occurred in a market, then both the monopolist and the entrant receive higher profits in that market if the monopolist acts cooperatively. Second, for each market the monopolist receives even higher profits if entry has not occurred. Third, for each market, if the monopolist is sure to act aggressively, then the potential entrant is better off not entering.

Intuition would suggest that if the above game were to actually be played, the monopolist would likely act aggressively whenever entry occurs in an early market, and in turn this would limit the number of markets in which entry occurs. However, the last period problem suggests that this intuition is incorrect. That is, because the monopolist will necessarily have an incentive to cheat (cooperate) in the last period, the unique equilibrium for the above game is that entry occurs in every market, and the monopolist cooperates each time entry occurs.[23]

KWMR present one way in which this paradox can be resolved. Their approach is to add uncertainty in terms of the payoff structure for the monopolist. In particular, they introduce a small probability that, if entry occurs in a market, the monopolist actually receives a higher payoff in that market by acting aggressively. Following Kreps and Wilson, when the monopolist has this new payoff structure we will refer to him as being strong, while when he has the old payoff structure we will refer to him as being weak. What KWMR demonstrate is that the introduction of this small probability the monopolist is strong has a dramatic impact on the nature of equilibrium. That is, when this small probability is present a typical equilibrium will be for no entry to occur until there are only a few markets remaining, rather than the zero probability equilibrium in which entry occurs in each market. The intuition for this result is that, even if the monopolist does not get a higher immediate return by responding aggressively to entry, entry in an early stage will still bring forth aggressive behavior because of the long term return. The long term return derives from the fact that by acting aggressively in response to an early entry, the weak monopolist causes later potential entrants to remain uncertain as to the monopolist's type.

As we stated earlier, the above result can also be understood in terms of responders versus nonresponders. In our terminology, what KWMR do is introduce a probability that the monopolist is a nonresponder. That is, they introduce a probability that the monopolist always acts aggressively.

Further, the environment they consider exhibits synergism. The logic is as follows. If the potential entrant has nothing but prior information concerning which type of monopolist he faces, then the presence of a positive probability that the monopolist always acts aggressively lowers the return to entry. In turn, this raises the incentive for a weak monopolist to act aggressively because by doing so it keeps the uncertainty of the potential entrant intact. That is, there is synergism in that the higher the probability the monopolist always acts aggressively, the larger is the return for the weak monopolist to act aggressively. Further, given that it is the strong monopolist who is the nonresponder and that this is an environment which exhibits synergism, Section II tells us that it is the strong monopolist who should be disproportionately important. The result, as found by KWMR, is that the introduction of just a small probability the monopolist is strong causes a dramatic change in the nature of the equilibrium.[24]

Overall, this section has focused on two different environments where responders and nonresponders are present because of differences in tastes across agents. The first environment was a labor market setting where workers have the opportunity to shirk, while the firm offers a non-market clearing wage in order to deter shirking. Because that environment exhibits congestion, we were able to show that it is the workers who receive disutility from effort, i.e., the responders, who are disproportionately important. The second environment was the entry deterrence setting previously analyzed by Kreps/Wilson and Milgrom/Roberts. That environment exhibits synergism. Hence, it is not surprising that it is the strong monopolist, i.e., the nonresponder, who is disproportionately important in that setting.

VI. Conclusion

Many economic environments exhibit the heterogeneity we refer to as responders versus nonresponders. That is, in many environments there will be one set of agents whose behavior depends on actions chosen by other agents in the population (responders), while for a second set of agents behavior is independent of the actions chosen by others (nonresponders). In this paper we have considered environments characterized by this heterogeneity, with the focus being on the following question. When both types of agents are present in an environment, is it the responders or is it the nonresponders who are disproportionately important?

We began the paper by presenting a general framework for answering the above question. To summarize, whether it is the responders or the nonresponders who are disproportionately important depends on the nature of the interaction among agents. When the return to an agent choosing an action decreases with the total number of agents who choose that action, i.e., when the environment exhibits congestion, then it is the responders who are disproportionately important. The reason is that, in this case, the presence of nonresponders is somewhat nullified by the "response" of the responders to their presence. On the other hand, when the return to an agent choosing an action is positively related to the total number of agents who choose that action, i.e., when the environment exhibits synergism, then it is the nonresponders who are disproportionately important. The logic here is that responders now "respond" in a manner which reinforces, rather than nullifies, the behavior of the nonresponders.

We then proceeded to both demonstrate and discuss how this approach can be used to help understand a variety of economic environments. Particular applications considered are: (i) the recent work on the evolution of market outcomes given network externalities; (ii) recent research on heterogeneity in information processing abilities; (iii) the pioneering work on reputation of Kreps/Wilson and Milgrom/Roberts; and (iv) a labor market setting where workers have the opportunity to shirk, while the firm offers a non-market clearing wage in order to deter shirking.

Appendix

Proof of Proposition 1: Let $\bar{c}$ and $\underline{c}$ be such that

$$LF_x(0)-\underline{c}=LF_y(\bar{N})-c_y \quad \text{(A1)}$$

and

$$LF_x(\bar{N})-\bar{c}=LF_y(0)-c_y. \quad \text{(A2)}$$

Given $F_x'>0$ and $F_y'>0$, it is clear that $\bar{c}>\underline{c}$. Further, it should be clear that if an agent is going to buy a unit of one of the products, then he will buy it at birth. Suppose $c_x<\underline{c}$ and consider agent i. (A1) tells us that even if all other agents (both at any date and across all dates) puchase product Y, then this consumer will still have an incentive to purchase X. Hence, the unique equilibrium is that all consumers purchase X.

Suppose $c_x>\bar{c}$ and consider agent i. (A2) tells us that even if all other agents purchase product X, then this consumer will still have an incentive to purchase Y. Hence, the unique equilibrium is that all consumers purchase Y.

Finally, suppose $\underline{c}\le c_x\le\bar{c}$ and consider agent i. If all other agents purchase Y, then (A1) tells us the agent is at least as well off purchasing Y as purchasing X. Similarly, if all other agents purchase X, then (A2) tells us the agent is at least as well off purchasing X as purchasing Y. Hence, it is an equilibrium for all agents to purchase Y, and it is an equilibrium for all agents to purchase X.

<u>Proof of Proposition 2</u>: Let $\underline{c}'$ and $\bar{c}'$ be such that

(A3) $$LF_x(\bar{N})-\bar{c}'=LF_y(0)-c_y$$

and

(A4) $$\int_0^L F_x(\bar{N}-tn(t))dt-\underline{c}'=\int_0^L F_y(tn(t))dt-c_y.$$

A comparison of (A2) and (A4) yields $\bar{c}'=\bar{c}$. A comparison of (A3) and (A4) yields $\bar{c}'>\underline{c}'$. A comparison of (A1) and (A4) yields $\underline{c}'>\underline{c}$.

Prior to date T^* only X is available, and the invention of Y is not anticipated by agents born before date T^*. Hence, given the restriction in footnote 9, all agents born prior to date T^* purchase X at birth. Suppose $c_x>\bar{c}'$ and consider agent i born T^* or later. (A2) tells us that even if all other agents purchase product X, then this agent will still have an incentive to purchase Y. Hence, the unique equilibrium is that starting at date T^* all agents purchase Y at birth.

Suppose $c_x<\underline{c}'$ and consider agent i born at date T^*. Further, suppose all other agents born T^* or later purchase Y at birth. (A3) tells us that agent i will still have an incentive to purchase X at birth. Recursively repeating this argument one can demonstrate that any agent born date T^* or later will have an incentive to purchase X at birth. Hence, the unique equilibrium is that starting at date T^* all agents purchase X at birth.

Finally, suppose $\underline{c}'\leq c_x\leq\bar{c}'$. Consider agent i born at date T^* or later. If all other agents purchase X at birth, then (A2) tells us that the agent is at least as well off purchasing X as purchasing Y. On the other hand, if starting at date T^* all other agents purchase Y at birth, then (A3) tells us that the agent is at least as well off purchasing Y as purchasing X. Hence, it is an equilibrium for all agents to purchase X, and it is also an equilibrium for all agents born T^* or later to purchase Y.

<u>Proof of Proposition 3</u>: $\hat{c}$ is defined by (A5).

$$LF_x(\bar{N})-\hat{c}=LF_y(\bar{N})-c_y \tag{A5}$$

Further, let $\hat{c}'$ be defined by (A6).

$$LF_x(\bar{N})-\hat{c}'=\int_0^L F_y(tn(t))dt-c_y \tag{A6}$$

A comparison of (A5) and (A6) yields $\hat{c}'>\hat{c}$. Further, a comparison of (A3), (A4), and (A6) yields $\underline{c}'<\hat{c}'<\bar{c}'$.

Using the same logic as in the proof of Proposition 2, all agents born prior to date T^* purchase X at birth. Suppose $c_x<\hat{c}'$. Given $\hat{c}'<\bar{c}'$, it is an equilibrium for all agents to purchase X. Further, (A6) tells us that agents born at date T^* will prefer this equilibrium over one where starting at date T^* only Y is purchased. Hence, all agents purchase X at birth.

Suppose $c_x>\hat{c}'$. Given $\hat{c}'>\underline{c}'$, it is an equilibrium for all agents born T^* or later to purchase Y. Further, (A6) tells us that agents born at date T^* will prefer this equilibrium over one where all agents purchase X. Hence, all agents born T^* or later purchase Y.

<u>Proof of Proposition 4</u>: Let x_i^* be the value for x_i in the initial equilibrium (the initial equilibrium must be symmetric), X^* be the value for X in the initial equilibrium, and x_i^m be the value for x_i for a maximizer. π^n and π^m are given respectively by

$$\pi^n=x_i^*-(\alpha^*+\epsilon)c(x_i^*)f(X) \tag{A7}$$

and

(A8) $$\pi^m = x_i^m - (\alpha^* + \epsilon) c(x_i^m) f(X).$$

The first order condition for the initial equilibrium is given by (A9), while the first order condition for x_i^m is given by (A10).

(A9) $$1 - \alpha^* c'(x_i^*) f(X^*) = 0$$

(A10) $$1 - (\alpha^* + \epsilon) c'(x_i^m) f(X) = 0$$

(A7) and (A8) yield (A11) and (A12). Note, the restriction in footnote 13 guarantees that X and x_i^m are differentiable with respect to ϵ.

(A11) $$\left.\frac{d\pi^n}{d\epsilon}\right|_{\epsilon=0} = -c(x_i^*) f(X^*) - \alpha^* c(x_i^*) \frac{df(X^*)}{dX} \left.\frac{dX}{d\epsilon}\right|_{\epsilon=0}$$

and

(A12) $$\left.\frac{d\pi^m}{d\epsilon}\right|_{\epsilon=0} = -c(x_i^*) f(X^*) - \alpha^* c(x_i^*) f'(X^*) \left.\frac{dX}{d\epsilon}\right|_{\epsilon=0} + \left.\frac{dx_i^m}{d\epsilon}\right|_{\epsilon=0} (1 - \alpha^* c'(x_i^*) f(X^*)).$$

(A9), (A11) and (A12) yield (2).

Taking the derivative of (A10) with respect to ϵ at $\epsilon=0$ yields

(A13) $$-c'(x_i^*) f(X^*) - \alpha^* c''(x_i^*) f(X^*) \left.\frac{dx_i^m}{d\epsilon}\right|_{\epsilon=0} - \alpha^* c'(x_i^*) f'(X^*) \left.\frac{dX}{d\epsilon}\right|_{\epsilon=0} = 0.$$

Let $R = (1/(\alpha^* c''(x_i^*) f(X^*) + \alpha^* c'(x_i^*) f'(X^*) pN))$. Since $\left.\frac{dX}{d\epsilon}\right|_{\epsilon=0} = pN \left.\frac{dx_i^m}{d\epsilon}\right|_{\epsilon=0}$, we can rearrange (A13) to yield

(A14) $$\left.\frac{dx_i^m}{d\epsilon}\right|_{\epsilon=0} = -c'(x_i^*) f(X^*) R < 0$$

or

(A15) $$\left.\frac{dX}{d\epsilon}\right|_{\epsilon=0} = -pNc'(x_i^*) f(X^*) R < 0.$$

Now take the derivative with respect to p.

(A16) $$\left.\frac{d^2X}{d\epsilon\, dp}\right|_{\epsilon=0} = -Nc'(x_i^*) f(X^*) R + Nc'(x_i^*) f(X^*) R^2 \alpha^* c'(x_i^*) f'(X^*) pN < 0$$

We can sign (A16) because $R\alpha^{*}c'(x_i^{*})f'(X^{*})pN<1$. In turn, (A16) implies $\left.\frac{dX^m}{d\epsilon}\right|_{\epsilon=0} < \left.\frac{dX}{d\epsilon}\right|_{\epsilon=0}$ or (3).

We also have

$$\text{(A17)} \qquad \left.\frac{dW}{d\epsilon}\right|_{\epsilon=0} = pN\left.\frac{d\pi^m}{d\epsilon}\right|_{\epsilon=0} + (1-p)N\left.\frac{d\pi^n}{d\epsilon}\right|_{\epsilon=0}.$$

Given (2) and (A11), this reduces to

$$\text{(A18)} \qquad \left.\frac{dW}{d\epsilon}\right|_{\epsilon=0} = -N[c(x_i^{*})f(X^{*})+\alpha^{*}c(x_i^{*})f'(X^{*})\left.\frac{dX}{d\epsilon}\right|_{\epsilon=0}]$$

or

$$\text{(A19)} \qquad \left.\frac{d(W-W^m)}{d\epsilon}\right|_{\epsilon=0} = -N\alpha^{*}c(x_i^{*})f'(X^{*})\left.\frac{d(X-X^m)}{d\epsilon}\right|_{\epsilon=0}.$$

(3) and (A19) yield (4).

<u>Proof of Proposition 5</u>: Let $x_i^{m,m}$ denote the behavior of a maximizer when all agents are maximizers. We can prove the first statement in the proposition by showing that if $f'>(<)0$, then $x_i^{m,m}<(>)x_i^m$. From (A10) we get that the first order conditions for x_i^m and $x_i^{m,m}$ can be written as

$$\text{(A20)} \qquad 1-(\alpha^{*}+\epsilon)c'(x_i^m)f(pNx_i^m+(1-p)Nx_i^{*})=0$$

and

$$\text{(A21)} \qquad 1-(\alpha^{*}+\epsilon)c'(x_i^{m,m})f(Nx_i^{m,m})=0.$$

Given the restriction in footnote 13, it must be the case that $x_i^{m,m}<x_i^{*}$. Hence, (A20) and (A21) tell us that if $f'>(<)0$, then $x_i^{m,m}<(>)x_i^m$.

Now consider (A18). Given $\left.\frac{dX^n}{d\epsilon}\right|_{\epsilon=0}=0$, we have

$$\text{(A22)} \qquad \left.\frac{dW^n}{d\epsilon}\right|_{\epsilon=0} = -N[c(x_i^{*})f(X^{*})],$$

$$\text{(A23)} \qquad \left.\frac{dW^m}{d\epsilon}\right|_{\epsilon=0} = -N[c(x_i^{*})f(X^{*})+\alpha^{*}c(x_i^{*})f'(X^{*})\left.\frac{dX^m}{d\epsilon}\right|_{\epsilon=0}],$$

and

$$\text{(A24)} \qquad \left.\frac{dW}{d\epsilon}\right|_{\epsilon=0} = -N[c(x_i^*)f(X^*)+\alpha^*c(x_i^*)f'(X^*)\left.\frac{dX}{d\epsilon}\right|_{\epsilon=0}].$$

Given $\left.\frac{dX^m}{d\epsilon}\right|_{\epsilon=0}<0$, (A22) and (A23) imply $\left.\frac{dW^m}{d\epsilon}\right|_{\epsilon=0}>(<)\left.\frac{dW^n}{d\epsilon}\right|_{\epsilon=0}$ if $f'>(<)0$.

The first statement of the proposition implies that if $f'>(<)0$, then $\left.\frac{dX}{d\epsilon}\right|_{\epsilon=0}<(>)p\left.\frac{dX^m}{d\epsilon}\right|_{\epsilon=0}$. We also know $\left.\frac{dX^m}{d\epsilon}\right|_{\epsilon=0}<0$. Hence, (A22), (A23) and (A24) imply that if $f'>(<)0$, then $\left.\frac{dW}{d\epsilon}\right|_{\epsilon=0}>(>)p\left.\frac{dW^m}{d\epsilon}\right|_{\epsilon=0}+(1-p)\left.\frac{dW^n}{d\epsilon}\right|_{\epsilon=0}$.

<u>Proof of Proposition 6</u>: Consider (A16). Taking the derivative of $\left.\frac{d^2X}{d\epsilon dp}\right|_{\epsilon=0}$ with respect to $f'(X^*)$ yields

$$\text{(A25)} \qquad 2Nc'(x_i^*)f(X^*)R^2\alpha^*c'(x_i^*)pN[1-R\alpha^*c'(x_i^*)f'(X^*)pN].$$

Since $R\alpha^*c'(x_i^*)f'(X^*)pN<1$, (A25) implies $\left.\frac{d^2X}{d\epsilon dp}\right|_{\epsilon=0}$ decreases as $f'(X^*)$ decreases. In turn, this implies $\left.\frac{d(X-X^m)}{d\epsilon}\right|_{\epsilon=0}$ increases as $f'(X^*)$ decreases.

Now consider (A19). If $f'<0$, then the previous result implies that a decrease in $f'(X^*)$ causes $\left.\frac{d(W-W^m)}{d\epsilon}\right|_{\epsilon=0}$ to increase.

<u>Proof of Proposition 7</u>: At $p=0$ it is obviously the case that $W=H$. Further, we impose restrictions on the model which guarantee that $W>H$ for any $p>0$ (see footnote 21). Given this, consider worker i who receives disutility from effort. If this worker is employed at the firm, he will (will not) shirk as long as $m_i<(>)m^*$, where

$$\text{(A26)} \qquad m^* = \frac{U(W)-U(W-k)}{U(W)-U(H-k)} .$$

Notice, $\frac{dm^*}{dW}<0$. Let $g(m)=G'(m)$ $(g(1)=0)$, $q=\lim_{m\to 1} 1-G(m)$, and let $\pi(p,W)$ denote profits per worker as a function of p and W. $\pi(p,W)$ is given by (A27).

(A27) $$\pi(p,W)=(1-p+pq)(X-W)+p(X-W)\int_{m^*}^{1} g(m)\,dm+p(Y-W)\int_0^{m^*}(1-m)g(m)\,dm$$

(A27) yields the following first order condition.

(A28) $$-(1-p+pq)-p\int_{m^*}^{1} g(m)\,dm-p\int_0^{m^*}(1-m)g(m)\,dm+pg(m^*)\frac{dm^*}{dW}[(1-m^*)(Y-W)-(X-W)]=0$$

Now consider p_1 and p_2, $p_1>p_2>0$, and let W_1 and W_2 be the respective wages which hold at each value for p. All we need to demonstrate is that $W_1>W_2$. We will first prove $W_1\neq W_2$. (A29) is the partial derivative of the left hand side of (A28) with respect to p.

(A29) $$1-q-\int_{m^*}^{1} g(m)\,dm-\int_0^{m^*}(1-m)g(m)\,dm+\frac{dm^*}{dW}\,g(m^*)[(1-m^*)(Y-W)-(X-W)]>0$$

Since the partial derivate is strictly greater than zero we have $W_1\neq W_2$. That is, if we plug p_2,W_2 into (A28), (A29) tells us that if we hold the wage fixed and increase p to p_1 then (A28) will no longer hold as an equality.

We will now demonstrate $W_1<W_2$. We know that W_1 maximizes profits per worker when $p=p_1$ and W_2 maximizes profits per worker when $p=p_2$. Hence, we have

(A30) $$\pi(p_2,W_2)-\pi(p_2,W_1)\geq 0$$

or

(A31) $$(1-p_2+p_2q)(W_1-W_2)+p_2B\geq 0,$$

and

(A32) $$\pi(p_1,W_2)-\pi(p_1,W_1)\leq 0$$

or

(A33) $$(1-p_1+p_1q)(W_1-W_2)+p_1B\leq 0,$$

where

$$B=\left\{(X-W_2)\int_{m^*(W_2)}^{1} g(m)\,dm+(Y-W_2)\int_0^{m^*(W_2)}(1-m)g(m)\,dm\right\}-\left\{(X-W_1)\int_{m^*(W_1)}^{1} g(m)\,dm+(Y-W_1)\int_0^{m^*(W_1)}(1-m)g(m)\,dm\right\}.$$

Suppose $W_1<W_2$. (A31) then yields B>0. In turn, given $p_1>p_2$, we now have

(A34) $$(1-p_1+p_1q)(W_1-W_2)+p_1B>(1-p_2+p_2q)(W_1-W_2)+p_2B.$$

However, (A34) contradicts (A31) and (A33). Hence, $W_1<W_2$.

<u>Proof of Proposition 8</u>: This is straightforward given Proposition 7. Let $\hat{W}$ be the wage offered when p=1. By definition $N_S=p\bar{N}\int_0^{m^*(W)} g(m)dm$ and $\hat{N}_S=\bar{N}\int_0^{m^*(\hat{W})} g(m)dm$. Proposition 7 states that for every p<1, $W<\hat{W}$. We know $\frac{dm^*}{dW}<0$ (see the proof of Proposition 7). This yields that for every 0<p<1, $\int_0^{m^*(W)} g(m)dm>\int_0^{m^*(\hat{W})} g(m)dm$. Hence, given the above definitions, we now have that for every 0<p<1, $N_S>p\hat{N}_S$. Further, the statement for N_{NS} now follows immediately.

Footnotes

[1]When we say the agent has an incentive to cheat in the last period, we simply mean he has an incentive to behave in a manner different from that which he is trying to establish a reputation for. In the environment considered by Kreps/Wilson and Milgrom/Roberts, this means the monopolist has an incentive to act in a cooperative manner if entry occurs in the last period (see Section V for more detail).

[2]The concepts we refer to as congestion and synergism have a variety of different names in the literature. Examples include decreasing and increasing returns (see Hirshleifer (1982, 1985) and Schelling (1978)), strategic substitutes and strategic complements (see Bulow et al (1985) and Cooper and John (1985)), and, as mentioned in the introduction, network externalities for the synergism case (see the references in the introduction). We employ the terms congestion and synergism because the literature has not settled down on a particular set of terms, and we would like to stay consistent with our own earlier work.

[3]If all responders were identical, then synergism implies that either all choose action A or all choose action B. Hence, in the argument above we are implicitly assuming that there is heterogeneity within the responder group.

[4]The main difference between the two models is that we consider an overlapping generations setting, while in their model agents are infinitely lived – although the agents do vary as regards the dates in which they are born. The overlapping generations specification is not at all crucial for the results which follow, but rather we simply found it more realistic.

[5]The assumption that the innovation is not anticipated by agents born

prior to date T^* is not crucial for what we want to show. See footnote 7.

[6]A fourth possibility is that the individual purchases both products during his lifetime, deriving benefits from each product at different points in time. In what follows we place a restriction on the model which rules this out as possible equilibrium behavior.

[7]Note, it turns out that for what we want to show neither this restriction nor the one mentioned in footnote 5 is critical. The reason is that, even without imposing these restrictions, agents born near date T^*-L will not want to purchase Y when it becomes available (this is true for any $c_y>0$). Hence, at least a subset of the agents born before date T^*, and who are alive at date T^*, are nonresponders. We impose the restrictions anyway for two reasons. First, given the restrictions, it is easier to identify who are the responders and who are the nonresponders. Second, imposing the restrictions makes our analysis more consistent with that of Farrell and Saloner.

[8]We also impose the restriction $LF_y(0)-c_y>0$. This eliminates the possibility of a degenerate equilibrium where nothing is purchased.

[9]In varying c_x we will now only consider the range $c_x<LF_x(0)$. If $c_x\geq LF_x(0)$ then the possibility arises that, for agents born prior to date T^*, no product is purchased.

[10]One might think that a similar result should hold for the comparison between $\bar{c}'$ and $\bar{c}$. The reason it doesn't is that both $\bar{c}'$ and $\bar{c}$ are determined by the value of c_x such that, when all other agents are buying X, an agent with both products available is indifferent between the two (see the proofs of Propositions 1 and 2).

[11]References to the work of Simon and his followers include Simon (1959), Cyert and March (1963), Williamson (1975), and Nelson and Winter (1982). See also Radner (1975) for explicit modeling of Simon's ideas.

[12]In Haltiwanger and Waldman (1986a) we considered a dynamic macroeconomic model exhibiting synergism wherein some agents satisfy a rational expectations assumption, while others satisfy an adaptive expectations assumption. In that model naive agents are only quasi-nonresponders. That is, in the period of a shock a naive agent's behavior is independent of the manner in which other agents adjust to the shock. However, in later periods such agents do respond to the behavior of other agents. The result is that naive agents are disproportionately important for the first few periods which follow a shock, while sophisticated agents may be disproportionately important in later periods.

[13]We also assume $c'(x)f(Nx)+c(x)Nf'(Nx)>0$ for all $x>0$. This rules out the possibility of multiple equilibria due to the presence of synergism.

[14]The assumption that the industry faces a perfectly elastic demand curve for its output is not at all critical for the results which follow. We impose it mainly to simplify the exposition. That is, employing a downward sloping demand curve would introduce an element of congestion into the model, and hence, the conditions translating into congestion and synergism would then be somewhat more complex.

[15]The structure of the model guarantees that all maximizers behave identically, and all nonmaximizers behave identically.

[16]Proposition 4 follows exactly Akerlof and Yellen's methodology for demonstrating that the private loss to being a nonmaximizer is second order, and that the presence of nonmaximizers has a first order impact on other

variables. For a discussion of this methodology see Akerlof and Yellen (1985a), p. 711.

[17]For discrete changes in α, statements concerning disproportionality and social welfare are difficult (see Haltiwanger and Waldman (1985), pp. 330-31, for a related discussion).

[18]Previous papers concerned with how the wage rate can deter shirking include Becker and Stigler (1974), Calvo and Wellisz (1979), Lazear (1981), Shapiro and Stiglitz (1984), and Yellen (1984).

[19]See Haltiwanger and Waldman (1986b) for other examples where responders and nonresponders are present because of differences in tastes across agents. In particular, that paper considers a number of examples where the heterogeneity arises because some agents are altruists, while others are egoists.

[20]We are assuming here that when the worker is self-employed he is better off expending effort than shirking.

[21]We impose two conditions sufficient to guarantee that for any $p>0$, the firm will want to offer a non-market clearing wage, i.e., $W>H$. The first is the above assumption that $G(m)$ has a mass point at $m=1$. The second is that $q(X-Y)>k$, where q is the probability weight on this mass point, i.e., $q=\lim_{m\to 1} 1-G(m)$.

[22]We would like to point out, however, that the result that workers who are willing to shirk are disproportionately important is not completely general. For example, we have considered the above model in a general equilibrium setting. Because of the resulting zero profit constraint, in that setting W <u>decreases</u> as p increases. Hence, that model exhibits synergism, and it is the workers who do not receive disutility from effort who are

disproportionately important. In that case, however, allowing the firm to have control over the probability of detection would introduce an element of congestion, with the overall result being that either type of agent could be disproportionately important.

[23]To be precise, the unique perfect Nash equilibrium for the game is that entry occurs in every market, and the monopolist cooperates each time entry occurs.

[24]Our description of Milgrom and Roberts above is somewhat imprecise. They actually show the paradox can be resolved by having the potential entrants only perceive there is a probability that the monopolist always acts aggressively. From our perspective this means that it is not actually necessary for nonresponders to be present for them to be disproportionately important. Rather, if the environment exhibits synergism, it may be enough for responders to simply perceive that nonresponders are present.

References

Akerlof, George A., and Janet L. Yellen(1985a), "Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?", American Economic Review, 75, September 1985, 708-720.

_______________(1985b), "The Macroeconomic Consequences of Near-Rational Rule-of-Thumb Behavior", Quarterly Journal of Economics, 100, Supplement 1985, 823-838.

Becker, Gary S., and George J. Stigler, "Law Enforcement, Malfeasance, and Compensation of Enforcers", Journal of Legal Studies, 3, 1974, 1-18.

Bulow, Jeremy I., Geanakoplos, John D., and Paul D. Klemperer, "Multimarket Oligopoly: Strategic Substitutes and Strategic Complements", Journal of Political Economy, 93, June 1985, 488-511.

Calvo, Guillermo A., and Stanislaw Wellisz, "Hierarchy, Ability and Income Distribution", Journal of Political Economy, 87, October 1979, 991-1010.

Conlisk, John, "Costly Optimizers Versus Cheap Imitators", Journal of Economic Behavior and Organization, 1, 1980, 275-293.

Cooper, Russell, and Andrew John, "Coordinating Coordination Failures in Keynesian Models", Cowles Foundation Discussion Paper No. 745, April 1985.

Cyert, Richard, and James March, A Behavioral Theory of the Firm, Englewood Cliffs: Prentice-Hall, 1963.

David, Paul A., "Clio and the Economics of QWERTY", American Economic Review, 75, May 1985, 332-337.

Farrell, Joseph, and Garth Saloner, "Standardization, Compatibility, and Innovation", Rand Journal of Economics, 16, Spring 1985, 70-83.

_______________, "Installed Base and Compatibility:

Innovation, Product Preannouncements and Predation", American Economic Review, 76, December 1986, 940-955.

Haltiwanger, John, and Michael Waldman, "Rational Expectations and the Limits of Rationality: An Analysis of Heterogeneity", American Economic Review, 75, June 1985, 326-340.

________________________________(1986a), "Limited Rationality and Synergism: The Implications for Macroeconomics", UCLA Working Paper No. 389, Revised April 1986.

________________________________(1986b), "The Role of Altruism in Economic Interaction", UCLA Working Paper No. 391, Revised June 1986.

Hirshleifer, Jack, "Evolutionary Models in Economics and Law: Cooperation Versus Conflict Strategies", Research in Law and Economics, 4, 1982, 1-60.

______________, "The Economic Approach to Conflict", UCLA Working Paper No. 320A, Revised May 1985.

Katz, Michael L., and Carl Shapiro, "Technology Adoption in the Presence of Network Externalities", Journal of Political Economy, 94, August 1986, 822-841.

Kreps, David, and Robert Wilson, "Reputation and Imperfect Information", Journal of Economic Theory, 27, August 1982, 253-279.

Layard, Richard, and Alan A. Walters, Microeconomic Theory, New York: McGraw Hill, 1978.

Lazear, Edward P., "Agency, Earnings Profiles, Productivity, and Hours Restrictions", American Economic Review, 71, September 1981, 606-620.

Milgrom, Paul, and John Roberts, "Predation, Reputation, and Entry Deterrence", Journal of Economic Theory, 27, August 1982, 280-312.

Nelson, Richard, and Sidney Winter, An Evolutionary Theory of Economic Capabilities and Behavior, Cambridge: Harvard University Press, 1982.

Radner, Roy, "Satisficing", Journal of Mathematical Economics, 2, June/September 1975, 253-262.

Russell, Thomas, and Richard Thaler, "The Relevance of Quasi Rationality in Competitive Markets", American Economic Review, 75, December 1985, 1071-1082.

Schelling, Thomas C., Micromotives and Macrobehavior, New York: Norton, 1978.

Selten, Reinhard, "The Chain-Store Paradox", Theory and Decision, 9, April 1978, 127-159.

Shapiro, Carl, and Joseph E. Stiglitz, "Equilibrium Unemployment as a Worker Discipline Device", American Economic Review, 74, June 1984, 433-444.

Simon, Herbert, "Theories of Decision-Making in Economics and Behavioral Science", American Economic Review, 49, June 1959, 253-283.

Williamson, Oliver, Markets and Hierarchies: Analysis and Antitrust Implications, New York: Free Press, 1975.

Yellen, Janet L., "Efficiency Wage Models of Unemployment", American Economic Review, 74, May 1984, 200-205.