Discussion Paper no. 911

OCCUPATIONAL CHOICE
AND THE PROCESS OF DEVELOPMENT

by

Abhijit V. Banerjee*

and

Andrew F. Newman**

July 14, 1990

Revised: October 27, 1990

*Department of Economics
Fisher Hall
Princeton University
Princeton, NJ 08544

**Department of Economics
Andersen Hall
Northwestern University
Evanston, IL 60201

We are grateful to Tim Besley, Allan Drazen, Kiminori Matsuyama, Joel Mokyr, Nancy Stokey and Aldo Rustichini for helpful discussion.

ABSTRACT

This paper models economic development as a process of institutional transformation by focusing on the interplay between agents's occupational decisions and the distribution of wealth. Because of capital market imperfections, poor agents choose working for a wage over self-employment, and wealthy agents become entrepreneurs who monitor workers. However, only with sufficient inequality will there be employment contracts; otherwise, depending on average wealth, there is either stagnation or self-employment. Thus, in a static context, the occupational structure depends on distribution. Since the distribution of wealth is itself endogenous, however, we demonstrate the robustness of this result by extending the model dynamically and studying examples in which initial wealth distributions have long run effects. In one case the economy develops into prosperity or stagnation, depending on the initial distribution; in the other example, it develops either widespread cottage industry (self-employment) or factory production (employment contracts).

OCCUPATIONAL CHOICE
AND THE PROCESS OF DEVELOPMENT

1. Introduction

Why does one country remain populated by small proprietors, artisans and peasants while another becomes a nation of entrepreneurs employing industrial workers in large factories? Why should two seemingly identical countries follow radically different development paths, one leading to prosperity, the other to stagnation? Questions like these are of central concern to both development economists and economic historians, who have been concerned with the study of the evolution of institutional forms, particularly those under which production and exchange are organized. Yet most of these institutional questions have resisted formal treatment except in a static context (see Stiglitz, 1988 for a review), while the dynamic issues which are peculiarly developmental have for the most part been restricted to the more narrow questions of output growth or technical change. This paper takes a first step in the direction of providing a dynamic account of institutional change by focusing on the evolution of occupational patterns, the contractual forms through which people exchange labor services.[1]

We focus on occupational choice because there are several strands of interaction between its dynamics and the process of development. Most obvious of these is the interaction with the distribution of wealth. Occupational choices helps determine the distribution of income and, over the long haul, the distribution of wealth. These in turn affect savings rates, the amount of risk-bearing, fertility rates and possibly the composition of demand and production in the economy. The link with the economy's growth rate and hence the process of development is then clear.

---

[1] Following what is becoming standard practice, we use the term "occupation" in this contractual sense, rather than to denote the particular productive activity in which an agent engages. For instance, a bricklayer and an accountant would be in the same occupation if each is an independent contractor or each works for a wage.

Equally important, as we have mentioned, is the connection which is suggested when one considers the process of development as one of institutional transformation as well as economic growth (Khan, 1989; Stiglitz, 1988; and Townsend, 1987). One of the most significant elements of the institutional structure of any economy is the dominant form of organization of production: it has "external" effects considerably beyond the efficiency (or the lack of it) of current production. Some of these effects may be politico-economic, but there are also purely economic effects; it has been argued, for example, that the introduction of the factory system in the early years of the Industrial Revolution left the technology unaffected and generated little efficiency gain initially.[2] But it seems very likely that in the long run this new form of organization of production helped to make possible the major innovations of the Industrial Revolution.

On the other side, there are also several ways in which the process of development affects the pattern of occupational choices. Development alters the demand and supply for different types of labor and hence affects wages and the pattern of allocation between different occupations. It alters the nature of risks and the possibilities for making innovations. It also changes the distribution of wealth and since people at different wealth levels typically have different incentives for taking up different occupations, this affects the distribution among the occupations.

The aim of this paper is to build a model of development which focuses directly on this interplay between the pattern of occupational choice and the process of development. The basic structure of interaction the model we present here is very simple. Because of capital market imperfections people can borrow only limited amounts. As a result, occupations which require high levels of investment are beyond the reach of poor people, who therefore choose to work for other, wealthier, employers. The wage rate and the pattern of occupational choice is now determined by the condition that the labor market must clear.[3] Depending on labor market conditions and on their wealth, other

---

[2]See for instance Cohen (1981) and Millward (1981). The factory system did, of course, benefit those who adopted it but that may have been a pure transfer stemming from more efficient monitoring.

[3]As a static model of occupational choice, this model is a simplified version of the model in Newman (1990), which also discusses the advantages of the

agents become self-employed in low-scale production or remain unemployed.[4]

The pattern of occupational choice which is therefore determined by the initial distribution of wealth; but the structure of occupational choice in turn determines how much people save and what risks they bear. These factors then give rise to the new distribution of wealth in the next period. It is the long run behavior of the dynamic process so generated that we study in this paper.

It turns out that despite the very simple structure of the model, our model does not fall into a category with which we are most familiar. As a rule, the dynamics are nonlinear and the state space — the set of all wealth distributions — very large, so that reasonably complicated behavior may be expected. While a complete mathematical analysis of the model is beyond the scope of the present paper, we confine our attention to two special cases which admit considerable dimensional reduction. These examples afford complete study and with them we are able to generate robust and natural instances of hysteresis or long run dependence on initial conditions.

These instances are important because the existence of some kind of hysteresis is an old and central idea in development economics. In particular, the very simple model we develop here already yields some patterns of dependence on initial conditions, which are consistent with what we know about the historical process of development.

The first example we look at is one in which the ultimate fate of the economy — prosperity or stagnation — depends in a crucial way on the initial distribution of wealth. If the economy initially has a high ratio of very poor people to very rich people, then the process of development runs out of steam and ends up in a situation of high unemployment and low wages (this may happen even when the initial per capita income is quite high, as long the

---

capital-market-imperfections approach over preference-based approaches such as that of Kihlstrom-Laffont (1979). See also the related work of MacDonald (1982) and Eswaran-Kotwal (1989).

[4]The unemployment here is of the purely "voluntary" sort, in the sense that given wealth and the going wage, an agent who is unemployed is no worse off than an identical agent who is working. Perhaps "subsistence" would be a more descriptive term.

distribution is sufficiently skewed). By contrast, if the economy initially has few very poor people (the per capita income can still be quite low), it will "take off" and converge to a high wage, low unemployment steady state.

That an economy's long-term prosperity may depend on initial conditions is a familiar idea in the development literature, and there have been some recent papers which capture different aspects of this phenomenon in a formal model.[5] Our paper differs from these in several important respects. First, most of these papers' results depend on technological increasing returns, generated either by production technology itself or by the presence of various kinds of spillovers in productivity or technical change; our model relies instead on a kind of "pecuniary" increasing returns stemming from imperfections in the capital market. Second, distribution tends not to play a causal role in this literature. A notable exception is Murphy-Shleifer-Vishny (1989a), but there the mechanism is the structure of demand for produced commodities rather than the occupational choice as mediated by the capital market; moreover, their model is static and therefore does not endogenize the distribution.

Third, and most important, none of these papers emphasize the endogeneity of economic institutions as part of the process of development. This distinction is highlighted by a second example we examine in which there appears a different kind of dependence on initial conditions. In this example it turns out that the economy might converge to a steady state in which there is only self-employment and small-scale production; alternatively, it may end up in a situation in which an active labor market and both large- and small-scale production prevail. Which of the two types of production organization eventually predominates once again depends on the initial distribution of wealth. Specifically, it turns out that an economy which starts with a large number of relatively poor people is more likely to develop wage employment and large-scale production than is an economy with few very poor people. This result is attractive because it provides a formalization for the classical view that despite the fact that capitalism is the more dynamic economic system, its initial emergence does depend on the existence of

---

[5] See for example Romer (1986); Galor and Zeira (1987); Lucas (1988); Murphy, Shleifer, Vishny (1989a, 1989b); and Matsuyama (1990).

a population of dispossessed whose best choice is to work for a wage.

In the next section of this paper we set up the basic model. The main results on the dynamics of occupational choices and the process of development are in Section 3. We conclude in Section 4 with a brief discussion of directions for further research.

2. The Model

2.1 *Tastes, Endowments and Decision Structure*

We consider a model with a large (a continuum) population of agents with identical preferences; the population in period t is described by a distribution function $G_t(w)$ which gives the measure of the population with wealth less than or equal to w.

Each agent lives for one period; at the beginning of that period he receives a bequest that his "father" has left him. He also has an endowment of 1 unit of labor. Labor services are not costlessly observable; we shall elaborate on this assumption below.

Having received his bequest, the agent has to make his occupational choice. To this end he might apply for a loan of a certain amount of money. If he receives the loan he decides whether or not he will renege on it. If he does, the lender in turn must decide whether to try to collect. Collection is costly, however, so that lenders who try collecting will want to recover the amount of the loan plus the collection cost and possibly a fine. We assume that a lender's decision is constrained by the fact that the borrower's net wealth can never be negative, i.e. the borrower has the option of declaring bankruptcy.

Once the amount of loan a person is getting and whether or not he will renege is decided, he makes the occupational choice which determines how he invests his labor and capital. After committing to an occupation, the income realization is determined and the agent makes his consumption and bequest choices. At the end of the period, he bestows the bequest to his son and passes from the scene.

An agent's utility is given by $B^{\beta}C^{\gamma} - Z$, where C is his consumption of the sole physical good in the economy, B is the bequest he leaves his son, and Z is the amount of labor he supplies; we assume that $\beta + \gamma = 1$. Notice that the bequest motive is of the "warm glow" variety (Andreoni, 1989); this yields a much more tractable problem than that generated by other specifications of bequest motives. Denote the income realization by Y; the indirect utility associated with these preferences then takes the form δY (where $\delta \equiv \beta^{\beta}\gamma^{\gamma}$), so that *agents are risk-neutral*.

2.2 *Technology*

There are three ways one can invest in production in this economy.

(a) A safe asset

This is best thought of as a bank which is outside this economy and which can borrow and lend at the fixed international interest rate $\hat{r}$ which is such that $(1-\gamma)\hat{r} < 1$.[6] Buying this asset thus amounts to lending to this bank. Selling this asset amounts to borrowing from the bank. Because of the problems in enforcing the debt contract, the bank is not willing to lend arbitrary amounts to anybody who asks for a loan. We will discuss the bank's lending policy below.

(b) A risky investment project.

This is best thought of as a physical investment like a farm or machine. The projects are indivisible and come in units of size I. Each investment project requires an initial investment of I units and a labor input of 1 unit to succeed. Any smaller investment or labor input will not generate any returns. If the project succeeds it generates a random return rI, where r is r_0 or r_1 with probabilities 1-q and q respectively ($r_0 < r_1$) and has mean $\bar{r}$. We assume that self-employment is feasible in the sense that it produces enough output to cover its labor cost, i.e.

$$I(\bar{r} - \hat{r}) > 1.$$

(c) Entrepreneurial production

This is the case when someone hires μ workers, each at a competitive wage v, to work for him. A worker undertakes a project involving an investment of I′ units and generating a random return r′. We assume that r′ takes on the values r'_0 and r'_1 with probabilities 1-q′ and q′.

---

[6] This assumption makes sure that the long run dynamics are sensible in the sense that people's wealth levels do not grow without bound.

The most natural interpretation of this production technology is that the projects individual workers are running are exactly the same as the projects being run by the self-employed. To facilitate this interpretation and to simplify the notation and exposition, we assume that $I' = I$ and that the mean of r' is the same as the mean of r (note that $q' \neq q$, however). The returns on each of the projects belonging to a single entrepreneur are perfectly correlated.

Under the suggested interpretation, the main difference between these two types of production is not so much in the technology but rather in the contracts under which output is distributed. In one, the worker is running it for himself, that is, he is the claimant on output, while in the other, the worker is running it for someone else. This distinction is important though, because the workers cannot guarantee that they will work; the person who invests in these projects has to monitor them to make sure they do. We assume that by putting in an effort level of 1, one person can monitor $\mu > 1$ separate projects. The monitoring activity is indivisible, and it is impossible to monitor a person who monitors others. Finally, we assume that entrepreneurial production is feasible in the sense that at the lowest possible wage rate (= 1, since at a lower wage the worker is better off unemployed) it is more profitable than self-employment, i.e.

$$\mu[I(\bar{r} - \hat{r}) - 1] \geq I(\bar{r} - \hat{r}).$$

2.3 *Markets*

We have already briefly described the market for loans; the rest of the discussion of this market must wait till later. The goods market in this economy is clearly trivial. The only other market in the economy is the labor market. The demand for labor comes from entrepreneurial production and the supply from individuals' occupational choices. We assume that this is a perfectly competitive market where the wage moves to equate supply and demand.

3. Analysis of the Model

3.1 *The Occupational Choices*

Given the assumptions we have made there are only four occupational options:

1. Being unemployed.

2. Being a worker and working in a risky project for someone else.
3. Being self-employed and running a risky project for oneself.
4. Being an entrepreneur, running μ projects by hiring labor and using one's own labor to monitor them.

There may be a question of how we rule out the other possibilities. The possibility of having an entrepreneur controlling more than μ projects is ruled out by our assumption that one cannot monitor someone who is himself supposed to monitor others. Being a part-time entrepreneur (sharing with someone else) is ruled out by the indivisible monitoring technology and in any case would not be attractive because of risk-neutrality. Raising capital, the other reason for combining with someone else, is ruled out by the same contract enforcement problems that exist between the bank and borrowers: one partner could as easily default on another partner as default on the bank. The same arguments rule out combining self-employment with any other activity.

3.2 *The Loan Market*

The key element to our model of the loan market is that we allow for the possibility of a borrower reneging on a debt. The story we have in mind is similar to that proposed by Kehoe and Levine (1989). Suppose a borrower puts all of his wealth w up as collateral and borrows an amount L. The lender cannot observe whether the borrower actually invests the borrowed amount in the project, so the borrower might instead choose to renege, perhaps fleeing to the neighboring county. If he does, he succeeds in escaping the lender's attempts at recovering the loan with probability $\pi > 0$, in which case he keeps L and loses his collateral w. If he is caught, he forfeits both L and w. The borrower will therefore renege whenever $\pi L > w$. (Assume that projects, once invested in, cannot be dismantled and consumed and that the lender is able to come around and claim his share of the output once it is determined; thus there is no issue of reneging after the investment is made.) Knowing this, the lender will only agree to make loans that satisfy $L \leq w/\pi$. *All loans that will actually be made in equilibrium will satisfy the above constraint and the borrowers will never choose to renege.*

The only two reasons to borrow in this model are to finance self-employment and to finance being an entrepreneur. The target levels of capital are therefore I and μI (we assume that wages are paid at the end of the period so there is no necessity to finance them). Now assume that someone with a wealth level $w < I$ wants to become self-employed and therefore needs to borrow an amount $I - w$. From above he will be able to borrow this amount if and only if $I - w \le w/\pi$ which requires that

$$w \ge \frac{\pi}{\pi+1} I.$$

The expression on the right-hand side of the above inequality gives the minimum wealth level necessary to qualify for a loan large enough to finance self-employment. Call this number w^*.

The smallest wealth needed to borrow enough to be an entrepreneur is derived by a parallel argument. We call this wealth level w^{**} and note that the expression for it is

$$w^{**} = \frac{\pi}{\pi+1} \mu I.$$

Since μ exceeds unity, w^{**} is always greater than w^*.

The model of the capital market we have chosen here yields a rather extreme version of increasing returns to wealth. In effect, it is not terribly different from the models in Bernanke and Gertler (1989, 1990), Sappington (1983), or the numerous discussions of credit markets in the development literature (see Bell, 1988 for a survey). As we shall see below, the present model is simple enough in some cases to allow reduction to a dynamical system on the two-dimensional simplex, a procedure which would be impossible with a more elaborate specification.

Notice that we are not allowing for any equity type assets based on the physical assets described above: nobody sells shares in the firms they set up. Since everybody is risk neutral, no one needs insurance. Equity is then purely a source of capital like debt and therefore suffers from the same problems of contract enforcement. Equity financing will therefore be available only if debt financing is available, and as a result nothing is lost by excluding equity markets.

3.3 *Market Equilibrium*

Recall that the distribution of wealth at the beginning of period t is denoted by $G_t(w)$. The (expected) returns to self-employment and unemployment are given by the parameters of the model; the wage v determines the returns to the other two occupations. The returns and the borrowing constraints determine the occupational choice made at each level of wealth. Integrating these choices with respect to $G_t(w)$ gives us the demand and the supply of labor. To find the equilibrium in each period, we need to find the level of wages which clears the labor market (we can assume that the goods market clears; as for the capital market, the interest rate has already been fixed at $\hat{r}$).

The demand for labor can come only from the entrepreneurs so that only the part of the population with $w \geq w^{**}$ need to be considered. These people will choose to be entrepreneurs as long as entrepreneurship is more rewarding than the other occupations i.e. as long as:

$$\delta[w\hat{r} + \mu I(\bar{r} - \hat{r}) - \mu v] - 1 \geq \delta[w\hat{r} + I(\bar{r} - \hat{r})] - 1$$

(it is better than self-employment) and

$$\delta[w\hat{r} + \mu I(\bar{r} - \hat{r}) - \mu v] - 1 \geq \delta[w\hat{r} + v] - 1$$

(it is preferable to wage work). That it is also preferred to unemployment is guaranteed by the first inequality along with our assumption that self-employment is feasible. Putting the two inequalities together, we get

$$\mu I(\bar{r} - \hat{r}) - \mu v \geq \mathrm{Max}\{I(\bar{r} - \hat{r}), v\},$$

or, more simply,

$$v \leq \frac{\mu - 1}{\mu} I(\bar{r} - \hat{r}).$$

As long as v satisfies this inequality, everyone who can afford to be an entrepreneur will want to be one. If this inequality is not satisfied, there will be no entrepreneurs; $\frac{\mu-1}{\mu}I(\bar{r}-\hat{r})$ is therefore the maximum equilibrium wage and we denote it by $\bar{v}$. The labor demand is thus derived to be

$$\begin{array}{ll} 0 & \text{if } v > \bar{v} \\ [0, \mu(1-G_t(w^{**}))] & \text{if } v = \bar{v} \\ \mu(1-G_t(w^{**})) & \text{if } v < \bar{v} \end{array}$$

The supply of labor is also easily determined:

$$\begin{array}{ll} 0 & \text{if } v < 1 \\ [0, G_t(w^*)] & \text{if } v = 1 \\ G_t(w^*) & \text{if } 1 < v \leq I(\bar{r} - \hat{r}) \\ 1 & \text{if } v \geq I(\bar{r} - \hat{r}) \end{array}$$

It is apparent that the equilibrium wage will be 1 if $G_t(w^*) > \mu[1-G_t(w^{**})]$ and $\bar{v}$ if $G_t(w^*) < \mu[1-G_t(w^{**})]$. The singular case in which $G_t(w^*) = \mu[1-G_t(w^{**})]$ gives rise to an indeterminate wage in $[1,\bar{v}]$. The fact that the wage generically assumes one of only two values and that it depends on no more information about the distribution $G(\cdot)$ than its value at w^* and w^{**} are the keys to the dimensional reduction which so simplifies our analysis below.

To summarize, the pattern of occupational choice that is generated in equilibrium is as follows:

1. Anybody with initial wealth less than w^* will be a worker unless wages are exactly 1, which occurs if labor supply is greater than or equal to labor demand at all feasible wages. Since, at this wage, workers are indifferent between working and not working, the labor market will clear by having some of the potential workers work while the others remain idle.

2. Anybody with initial wealth between w^* and w^{**} will become self-employed. This follows from the fact, noted above, that someone who chooses to be a worker even when he could be self-employed, would do so only if $v \geq I(\bar{r} - \hat{r})$. But these wages cannot be sustained in equilibrium.

3. Anybody who starts with initial wealth above w^{**} will be an entrepreneur as long as $v < \bar{v}$. If $v = \bar{v}$, since all the potential entrepreneurs are equally happy to be self-employed instead, in equilibrium a fraction $1 - G_t(w^*)/\mu G_t(w^{**})$ of them will opt to be self-employed so that the labor market may clear.

Thus, despite the fact that everybody has the same abilities and the same preferences, different people choose different occupations. What is more, the occupational choices made by individuals depends on the overall distribution of wealth. For example, if everyone is above w^*, everyone will be self-employed. Employment contracts emerge only if some people are below w^* while others are above w^{**}. And with everyone below w^*, unemployment becomes the only option. Thus, *the institutional structure of the economy, as represented by the pattern of occupational choice, depends on the distribution of wealth.*[7] These results parallel those in Newman (1990). The question, of course, is whether the dependence of institutional structure which obtains in the short run also obtains in the long, when the distribution itself is endogenous.

---

[7]So does static efficiency. In this model, a first-best optimum is achieved only when everyone is self-employed; even though the employment contract is optimal from the point of view of the parties involved, an equilibrium with employment contracts cannot be first-best efficient (some resources are being spent on monitoring instead of direct production).

3.4 *The Dynamics of Lineage Wealth*

We have described how, given an initial distribution of wealth, the equilibrium wage and occupational choices are determined. Knowledge of the realization of project returns then gives us each person's income and bequests, from which we can calculate a starting wealth distribution for the next period.

Given the form of the utility function it is evident that people leave a fraction $1-\gamma$ of their realized income as bequest. It is then straightforward to derive the expression for the evolution of wealth for each choice of occupation. These are:

1. Unemployed: $w_{t+1} = (1-\gamma)\hat{r}w_t$.

2. Wage laborer: $w_{t+1} = (1-\gamma)[\hat{r}w_t + v]$.

3. Self-employed: $w_{t+1} = (1-\gamma)[w_t\hat{r} + I(r-\hat{r})]$, where r, of course, is a random variable. The distribution of w_{t+1} can be derived straightforwardly using this equation and the distribution of r.

4. Entrepreneur: $w_{t+1} = (1-\gamma)[w_t\hat{r} + \mu[I(r-\hat{r})-v]$. Once again, this is random and its distribution can be derived from the distribution of r.

The simple transition diagram in Figure 1 can be used to represent the dynamics of lineage wealth. The horizontal axis represents current wealth while the vertical axis shows next period's wealth. Assuming $v = \bar{v}$, everybody with wealth between 0 and w^* will choose to be a worker and their offspring's wealth as a function of their current wealth will be given by the line segment AB. As we have said above everybody between w^* and w^{**} will choose to be self-employed and their wealth dynamics are given by the two parallel lines CD and C′D′, each representing one realization of the random variable r. Since the wage is $\bar{v}$, everyone above w^{**} will either be an entrepreneur or be self-employed; the two parallel lines DE and D′E′ represent the dynamics for a self-employed person while the two parallel lines FG and F′G′ represent that for an entrepreneur.

A similar diagram can be constructed for the case where $v = 1$. The specific positions of the different lines in these diagrams depend, of course,

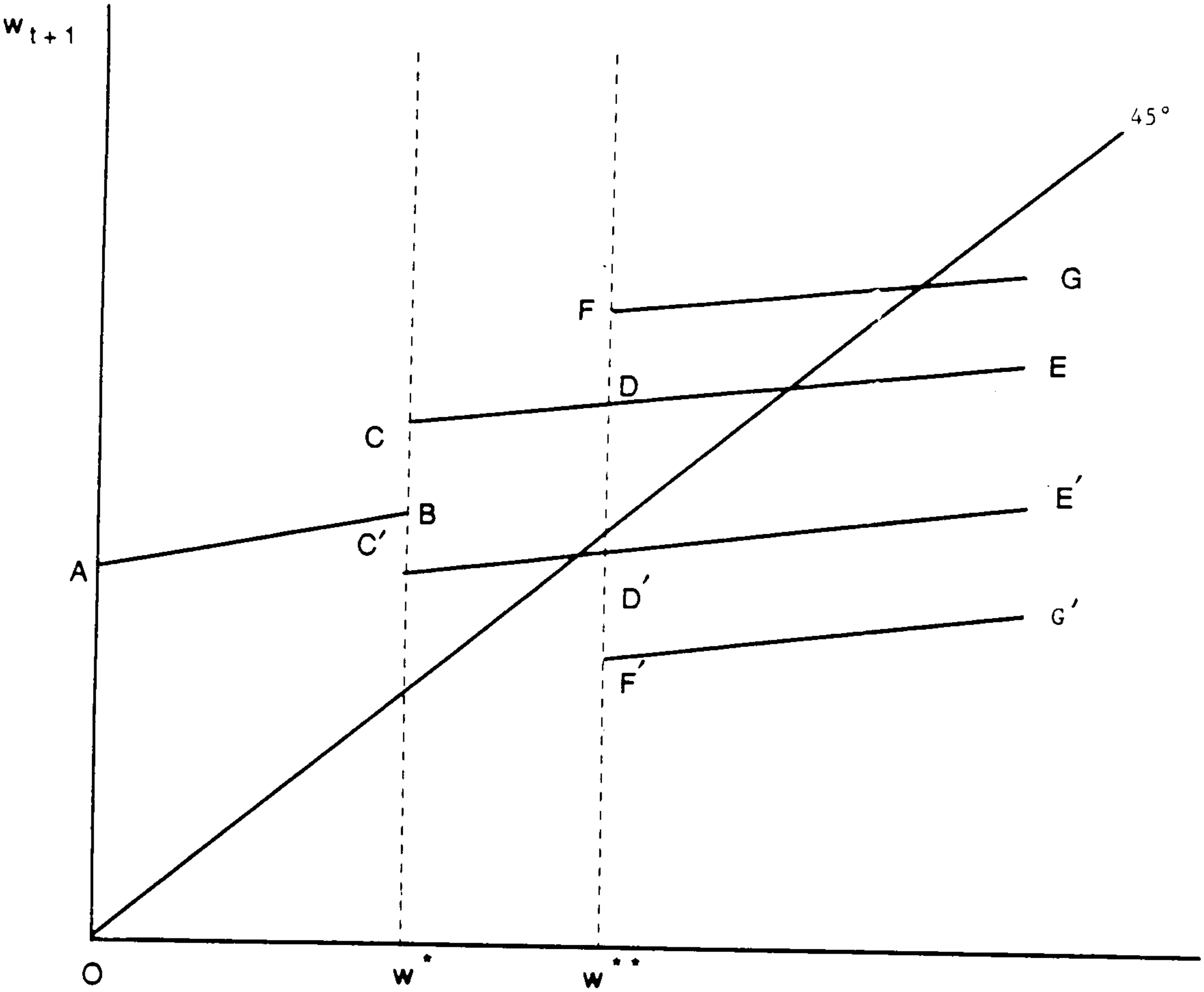

Figure 1 : $v = \bar{v}$

AB: $w_{t+1} = (1-\gamma)[w_t\hat{r} + \bar{v}]$

CE: $w_{t+1} = (1-\gamma)[w_t\hat{r} + I(r_1-\hat{r})]$

C′E′: $w_{t+1} = (1-\gamma)[w_t\hat{r} + I(r_0-\hat{r})]$

FG: $w_{t+1} = (1-\gamma)[w_t\hat{r} + \mu[I(r_1'-\hat{r})-\bar{v}]]$

F′G′: $w_{t+1} = (1-\gamma)[w_t\hat{r} + \mu[I(r_0'-\hat{r})-\bar{v}]]$

on the parameters of the model.

3.5 *The Dynamics of Distribution and Occupational Choice*

From the point of view of an individual lineage, wealth follows a Markov process. If this process were stationary, we could go ahead and use the standard techniques (see e.g. Stokey and Lucas, 1989) to establish existence and global stability of an ergodic measure on the wealth space and, since we are assuming a continuum of agents, reinterpret this to be the limiting wealth distribution for the economy. Under the stationarity assumption, one can study the Markov process by considering a (deterministic) map from the space of distributions to itself called a Markov operator (in the finite case this is just the familiar stochastic matrix). This operator is well known to be linear.

In our model, however, the stationarity assumption is not justified. Each lineage's transition rule depends in period t on the wage prevailing at that time. The wage in turn depends on the current distribution of wealth across all agents in the economy; as the distribution changes over time, so does the wage, thereby destroying the stationarity of the process.

In short, the state space for our model is not simply the wealth interval, but the set of distributions on that interval: this is the smallest set which provides us with all the information we need to fully describe the economy in period t and predict its characteristics in period t+1.

Now we have already shown that given the current distribution of wealth, we can determine the equilibrium level of wages and the pattern of occupational choices. Then, using the above transition equations, the current distribution of wealth $G_t(\cdot)$, and the fact that we have a large number of agents receiving independent project returns, we can (in principle) derive in a deterministic manner the distribution of wealth $G_{t+1}(\cdot)$ for the next period. We therefore have a well-defined, deterministic, transition map or operator on the space of wealth distributions, sending the current distribution of wealth to that of one period in the future.

Ordinarily, the operator so derived may be quite complex, and unlike the familiar Markov operator which is defined on the same space, it is nonlinear. The nonlinearity already tells us that uniqueness, global stability and other nice, easy-to-verify properties of linear systems are unlikely to obtain. But it would be desirable to be able to characterize the behavior of the economy somewhat more precisely than to simply say that it might, abstractly, display hysteresis, nonuniqueness, cyles or other nonlinear behavior.

The great attraction of the present model is that if we restrict attention to certain configurations of parameter values, we can achieve a rather precise characterization of the economy's behavior using elementary methods. In the rest of this section we will look at two examples which obtain when the individual transition diagrams like Figure 1 have certain specific configurations; these cases are illustrative of interesting historical patterns of development and occupational structure.

3.51 *Prosperity and Stagnation*

Consider the case when the transition diagrams for $v = 1$ and $v = \bar{v}$ are given by Figures 2(a) and 2(b). The configuration represented in these diagrams will obtain when w^* and $\bar{v}$ are relatively high, $1-\gamma$ is relatively low and the riskiness of the random realizations, given by the width of the bands, is quite large.

Look now at Figure 2(a). Defining $\bar{w}$ to be the fixed point of the intertemporal map $w_{t+1} = (1-\gamma)[w_t\hat{r} + \mu[I(r-\hat{r})-1]$, observe that this is the highest possible wealth level that can be sustained in the long run (any lineage with wealth greater than this value is sure to fall below it eventually). Without loss of generality then, we restrict all of our attention to wealth distributions on the interval $[0,\bar{w}]$.

Observe now that in Figure 2(a), a lineage currently with wealth in $[0,w^*)$ remains in that range in the next period. Anyone initially in $[w^*,w^{**})$ goes either to $[w^{**},\bar{w}]$ (if the project return is high) or to $[0,w^*)$ (if the project return is low). Finally, regardless of whether an agent who is initially in $[w^{**},\bar{w}]$ becomes self-employed or an entrepreneur, his offspring either remains there (if lucky) or goes to $[0,w^*)$ (if unlucky). The important point is that these transitions depend only on what interval one is in and not on the precise wealth level within that interval. Similarly, inspection of Figure 2(b) shows that when the prevailing wage is $\bar{v}$, the transitions between the same three intervals also depend only on those intervals and not on the

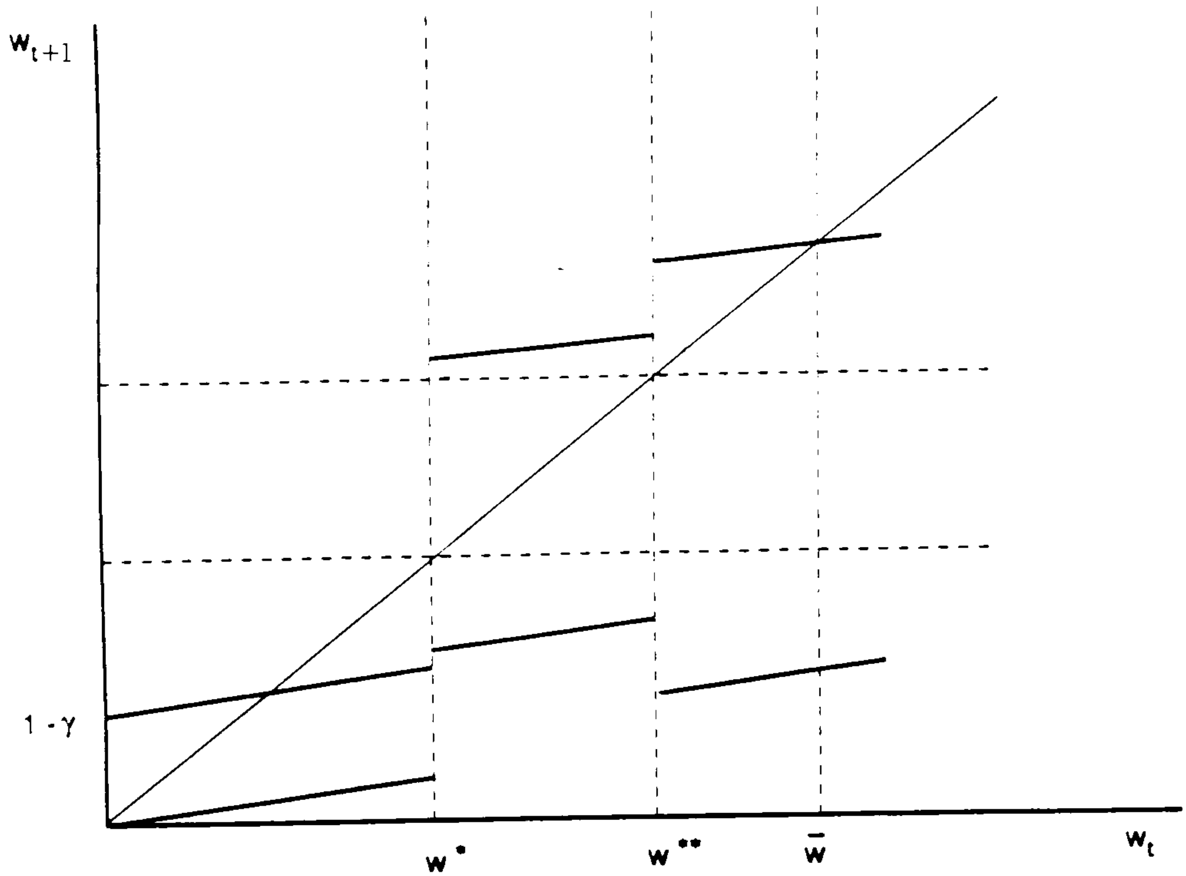

Figure 2(a): $v = 1$

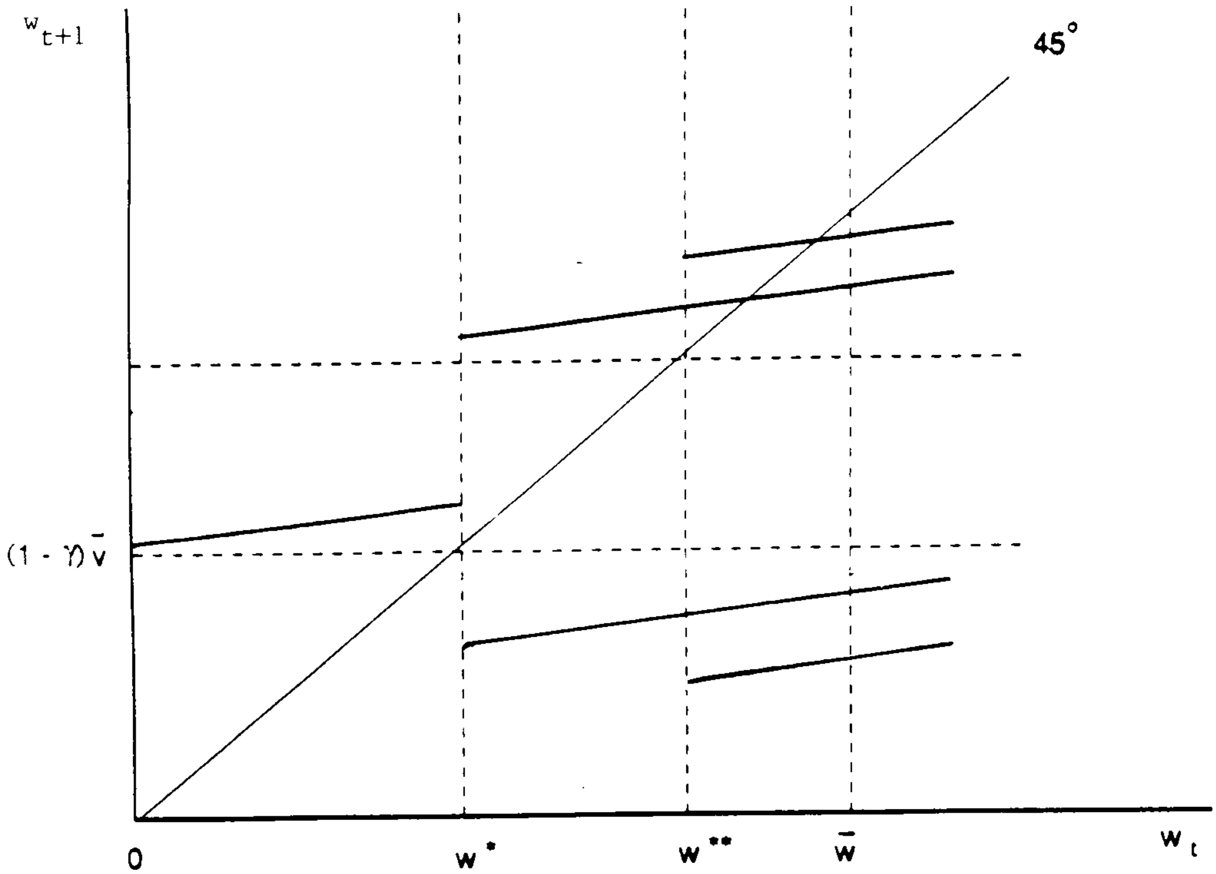

Figure 2(b): $v = \bar{v}$

wealth levels within them.[8]

As we showed in Section 3.3, the equilibrium wage and the occupational structure depend only on the ratio of the number of people in $[0,w^*)$ and the number of people in $[w^{**},\bar{w}]$, and not on any other properties of the distribution. Identify the three intervals with three "classes" L, M, and U (for lower, middle and upper); wealth distributions (fractions of the population in the three classes) are then given by probability vectors $p = (p_L,p_M,p_U)$, that is, points in Δ^2, the two-dimensional unit simplex. The state space for our economy is then just this simplex: for our purposes, it contains all the information we need.[9]

The evolution of the wealth distribution can now be represented by a dynamical system on Δ^2 which may be written

$$p_{t+1} = A(p_t)p_t, \tag{1}$$

where $A(p_t)$ is a 3×3 stochastic matrix[10] which depends on the current distribution p_t. The matrix assumes one of two forms: if $p_L > \mu p_U$, so that there is excess supply in the labor market and $v = 1$, then we have

$$A(p) = \begin{bmatrix} 1 & 1\text{-}q & 1\text{-}q' \\ 0 & 0 & 0 \\ 0 & q & q' \end{bmatrix}, \quad p_L > \mu p_U. \tag{2}$$

For the case of excess demand, the situation is slightly more complicated, since the individual transition probabilities for members of the class U depend on their occupation :

---

[8] The fact that w^* and especially w^{**} do not depend on the wage v is crucial to this simplification.

[9] In terms of our earlier notation, if $G(\cdot)$ is the current wealth distribution, then $p_L = G(w^*)$, $p_M = G(w^{**})-G(w^*)$ and $p_U = 1-G(w^{**})$. Of course, some information is lost by our dimensional reduction. For instance, if $H(\cdot)$ is another distribution with $H(w^*) = G(w^*)$ and $H(w^{**}) = G(w^{**})$, then it will be indistinguishable from $G(\cdot)$, even if, for instance, the two distributions have different means. The limiting distributions to which they converge will generally differ as well, but will be equal at w^* and w^{**}.

[10] The fact that $A(\cdot)$ is a stochastic matrix reflects only the fact that it maps Δ^2 to itself and should not be construed to mean that there is anything random about the process sending wealth distributions today into distributions tomorrow.

$$A(p) = \begin{bmatrix} 0 & 1-q & \begin{matrix}(1-q')(p_L/\mu p_U)+\\(1-q)(1-p_L/\mu p_U)\end{matrix} \\ 1 & 0 & 0 \\ 0 & q & \begin{matrix}q'(p_L/\mu p_U)+\\q(1-p_L/\mu p_U)\end{matrix} \end{bmatrix}, \quad p_L < \mu p_U. \tag{3}$$

These matrices are easily derived by inspecting Figure 2 and using the fact that the large number of agents lets us identify the probability of different project outcomes with the fractions of the population receiving those outcomes. For instance, when $v = \bar{v}$, $p_L/\mu p_U$ of the agents with wealth greater than w^{**} become entrepreneurs; of them, q' get the high return and remain with wealth greater than w^{**}, while $1-q'$ end up below w^*. The remaining agents in U become self-employed and go to L and U in the proportions $1-q$ and q. This gives us the lower right entry of the excess demand version of the matrix, and the other entries are derived similarly.

Now it will be convenient to study the dynamics of our economy by using a phase digram; to do so we take the continuous time analog of equation (1) and restrict our attention to the two variables p_L and p_U, since knowledge of these gives us p_M. This procedure gives us a piecewise-linear system of differential equations:

$$\dot{p}_L = \begin{cases} 1-q-(1-q)p_L+(q-q')p_U, & p_L > \mu p_U \\ 1-q-(2-q+q'/\mu-q/\mu)p_L, & p_L < \mu p_U \end{cases} \tag{4}$$

$$\dot{p}_U = \begin{cases} q-qp_L+(q'-q-1)p_U, & p_L > \mu p_U \\ q-(q+q/\mu-q'/\mu)p_L-p_U, & p_L < \mu p_U \end{cases} \tag{5}$$

The corresponding phase diagram is provided in Figure 3. The heavy line is the "boundary" $p_L = \mu p_U$ between the two linear systems.[11]

There are two stationary values distributions in this diagram, labelled S and P, and it is clear that they are both locally stable, with fairly large basins of attraction. But these stationary distributions are very different

---

[11] We have assumed that on the boundary the high-wage dynamics apply. The behavior at the boundary are of course affected by which wage is supposed to prevail there. Making alternative assumptions will not significantly change our results.

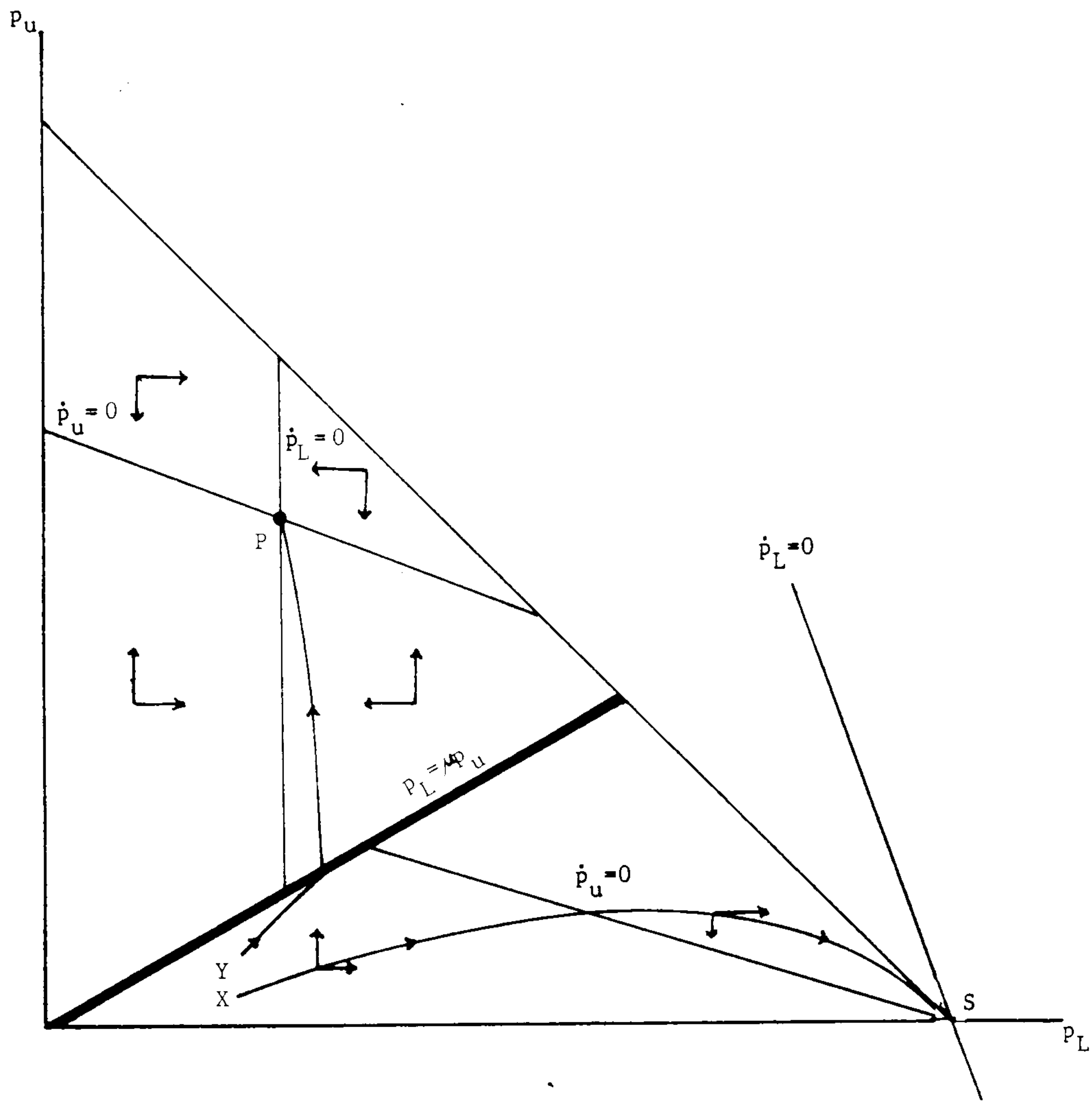

Figure 3

from each other. S is a state of economic collapse or stagnation: $p_L = 1$, so all agents have low wealth, which entails that they are all unemployed. By contrast, P is a prosperous economy with both self-employment and an active labor market in which workers receive high wages; since the transition probabilities between the states are relatively high, there is also considerable social mobility.

What then will be the long-run behavior of the economy? Clearly it depends on the initial conditions: *the long-term prosperity and occupational structure of the economy depends on the initial distribution of wealth.* Roughly speaking, economies where the initial ratio of workers to entrepreneurs is low are more likely to fall above the boundary line where they will be subject to the high-wage dynamics and are therefore more likely to converge to P. Where the initial ratio of poor to wealthy is high, the economy will be subject instead to the low-wage dynamics.

Does this mean that an economy in which the ratio of poor to wealthy is high is doomed to stagnate? One of the advantages of using the phase diagrams is that it lets us trace out the development path of the economy explicitly so that we can answer such questions. By examining Figure 3, we can see that the answer is no, particularly if the middle class is sufficiently large (distributions with a large middle class are located near the origin). Consider the path starting at the point Y. Here most agents in the economy are self-employed, while the few workers that there are receive low wages because there are so few entrepreneurs demanding their labor (recall that some agents in state L must be unemployed). Over time, some of the self-employed become entrepreneurs and the rest fall into the lower wealth class. Along this particular path, the number of agents in U grows sufficiently fast that all agents in L are eventually hired as workers, and the economy is brought to the boundary: now there is excess demand for labor and the high-wage dynamics take over, with the number wealthy agents growing rapidly (the number of workers declines slightly along this part of the development path, from which we infer that the ranks of the self-employed must be growing. Thus even though this economy begins with a high ratio of poor to wealthy, it eventually achieves prosperity.

What is interesting is that there may be initial conditions which are quite close to each other but which lead to very different long-run behavior. For instance, if instead of Y, we start at start at a point like X, the upper class grows slightly faster than the lower class, with both growing at the expense of the middle class of self-employed. The wage remains low, however, and eventually the lower class begins to dominate until the economy collapses to the stationary point S. The contrasting development paths that our model generates may shed light on the starkly different developmental histories of such countries as the United States and Argentina, which at the turn of the century both had large, affluent middle classes. Since then, of course, only the U.S. has developed into the mobile, prosperous economy such as that represented by P in our model.

Notice as well that because the model lets us study dynamic paths explicitly, we can check whether an economy might follow the Kuznets curve, or some other favorite developmental story. In the present example, it should be clear that such a curve is hardly an inevitable path to prosperity: although the path starting at Y does display this tendency (with the equality, as measured by the relative size of the middle class, first declining and then increasing), there are plenty of other paths which are more nearly monotonic.

We should point out that the picture given in Figure 3 is not the only possible depiction of equations (4) and (5). If the parameters μ, q and q′ are varied somewhat, the details of the phase diagram will vary as well. Behavior near the boundary between the high- and low-wage regions is most likely to be affected, but generally speaking, there will continue to be two locally stable stationary distributions with rather different characteristics. Their domains of attraction may vary, leading to somewhat different development paths, but otherwise, the behavior we have outlined is reasonably robust.[12]

3.52 *The Cottage vs. the Factory*

A somewhat different set of development paths can be generated with a different configuration of parameter values. Consider the case when the

[12] There is one exception to this rule which is at least possible to generate, although not particularly likely. If q, q′ and μ satisfy $\mu q(1-q) < 1+q'+q(q-q')$, it turns out that the stationary point to the high-wage dynamics will actually lie *below* the $p_L = \mu p_U$ boundary. In this instance, there is no longer any hysteresis, since in converging to the high-wage stationary point, the economy crosses the boundary and the low-wage dynamics take over; the economy inevitably stagnates.

transition map is as in Figure 4(a) and 4(b) (corresponding once again to the cases $v = 1$ and $v = \bar{v}$). It is easy to check that such a configuration will indeed arise for parameter values within a certain range. As in the previous section, the aggregate dynamic behavior can be reduced to a two-dimensional dynamical system in the simplex, as long as the initial wealth distribution is supported in the interval $[0,\bar{w}]$. Using the same definitions for the states as above, we follow a similar procedure to derive the transition matrices and to pass to continuous-time. We arrive at the following system of piecewise-linear differential equations:

$$\dot{p}_L = \begin{cases} 0, & p_L > \mu p_U \\ ((1-q')/\mu-1)p_L, & p_L < \mu p_U \end{cases} \tag{6}$$

$$\dot{p}_U = \begin{cases} q-qp_L+(q'-q-1)p_U, & p_L > \mu p_U \\ q-(q+q/\mu-q'/\mu)p_L-p_U, & p_L < \mu p_U \end{cases} \tag{7}$$

The phase diagrams for this set of differential equations is given in figure 5(a). The upper triangle represents the case when $v = \bar{v}$ and the lower triangle represents the case when $v = 1$. The heavy line is the "boundary" $p_L = \mu p_U$ between the two linear systems.

In the upper triangle the point C represents a stationary distribution and it is clearly locally stable. In the lower triangle there is whole range of stationary distributions since the $\dot{p}_L = 0$ locus is degenerate. This is a consequence of the fact that there is no way in or out of state L. Hysteresis of a specific type is therefore built into this model.

Since our interest is in hysteresis generated by workings of the labor market, we feel that it is best to eliminate this other type of hysteresis from the model. This is legitimate since all we need to get rid of it is to perturb the dynamics slightly, allowing very small probabilities of moving from state L to the other two states and from the other two states to L.[13] The phase diagram for these perturbed dynamics is given in diagram 5(b). Notice that, as expected, the $\dot{p}_U = 0$ loci in both triangles have moved only very slightly, as has the $\dot{p}_L = 0$ locus in the upper triangle. The most significant

[13]Think of these small probabilities as corresponding to winning the lottery and having a thunderbolt hit your house and factory.

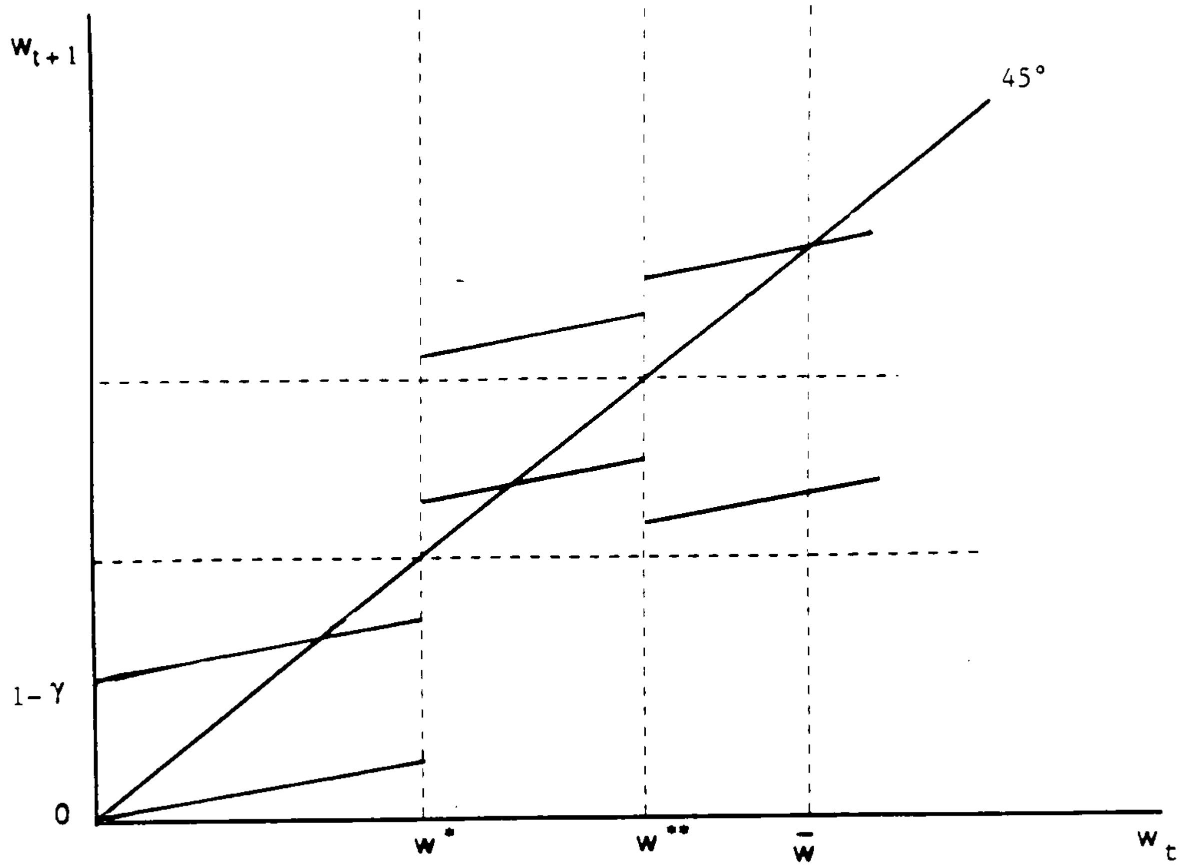

Figure 4(a): $v = 1$

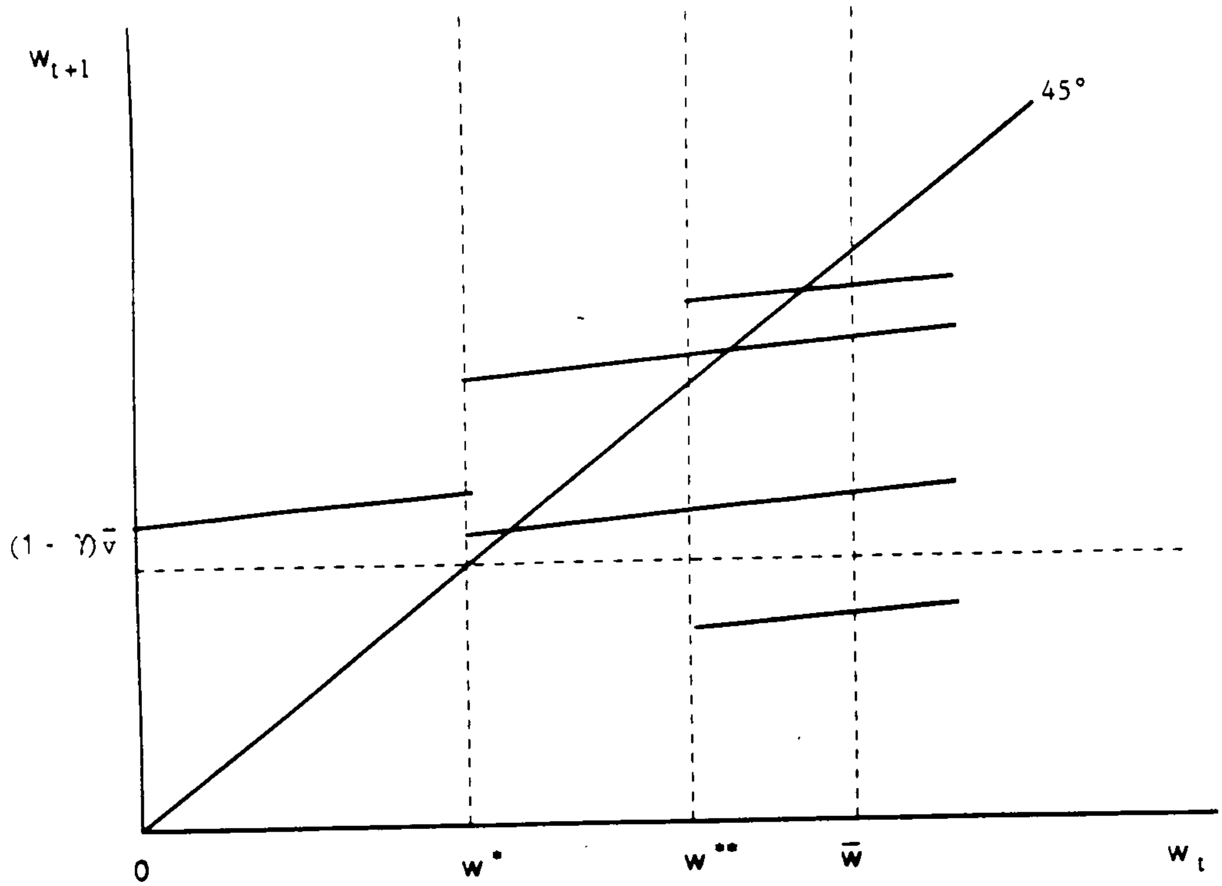

Figure 4(b): $v = \bar{v}$

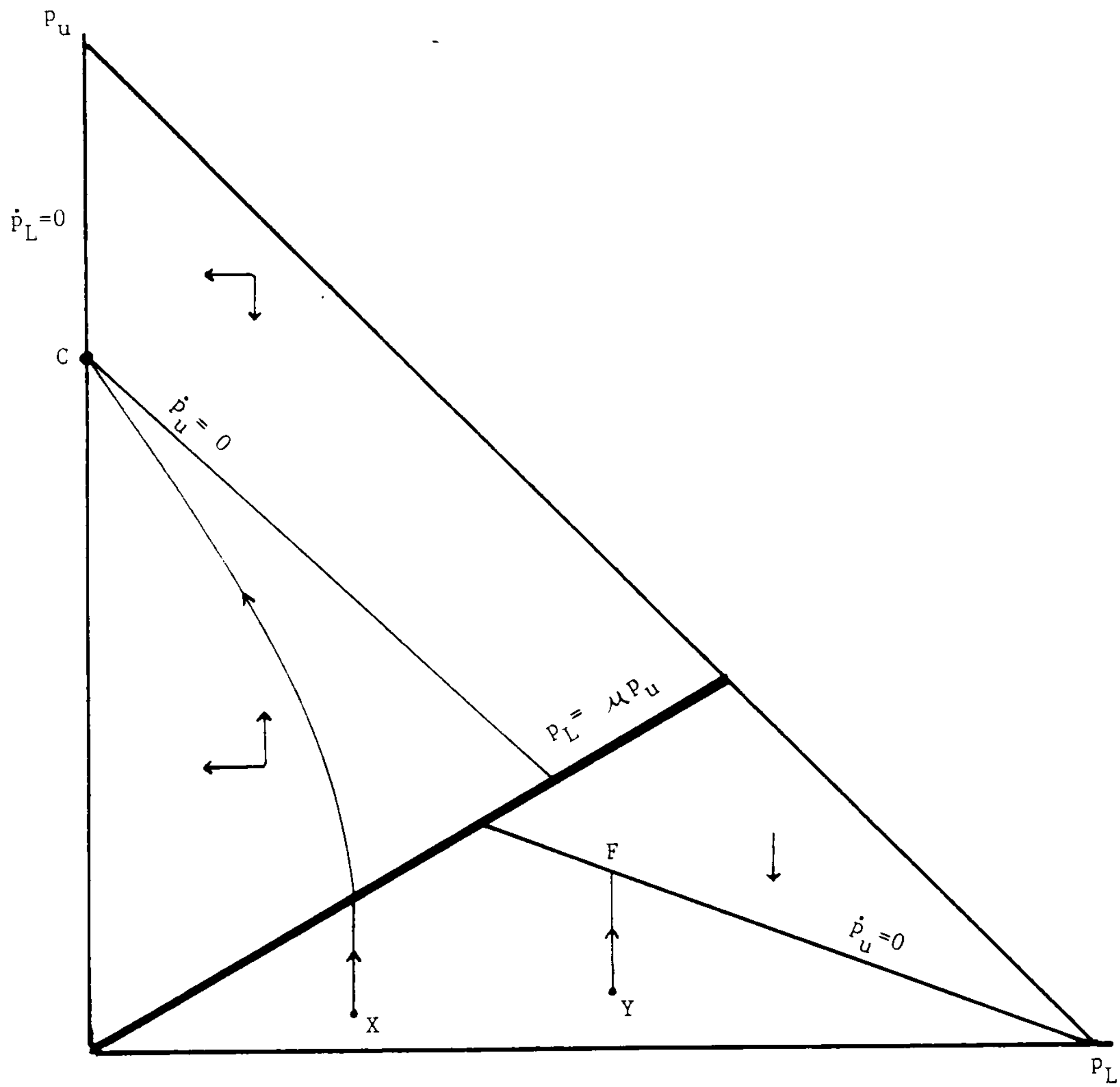

Figure 5(a)

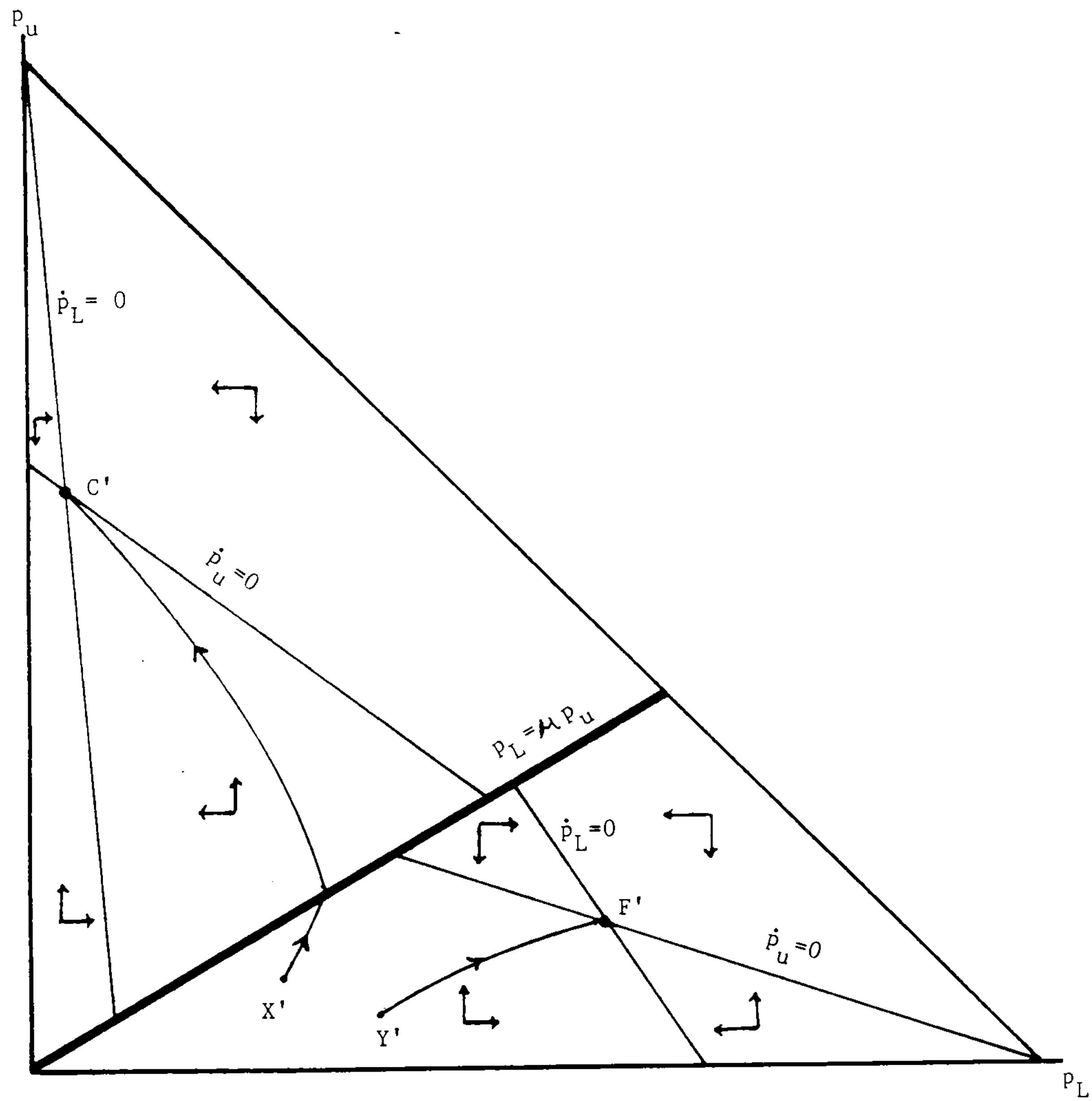

Figure 5(b)

change is that now we have a $\dot{p}_L = 0$ locus in the lower triangle which intersects the $\dot{p}_U = 0$ locus in that triangle at the point F′.

The point F′ represents a stationary distribution as does the point C′. Both of these points are clearly locally stable. However they also quite obviously represent very different social situations. F′ is an economy where there are three distinct classes with relatively little social mobility between the top two classes and the lowest class (recall that all the mobility in and out of the lowest class come from the small probabilities we have introduced to eliminate the degenerate hysteresis). The principle reason behind the limited social mobility is that the ratio of workers to entrepreneurs is high; the consequent low wage rate makes it impossible, given the propensity to save, for workers to accumulate enough wealth to enter state M. At the same time, the project returns (in particular the low ones) are high enough to insure the self-employed and entrepreneurs from against falling into state L.

C′, by contrast, is a situation in which there is really only one occupation in the economy; the overwhelming majority of the population is self-employed (in the unperturbed version of the model, *everyone* is self-employed). While the diagram shows that there are a substantial number of people in the class U, most of of these people are actually self-employed: although wealthy enough to be entrepreneurs, there are virtually no workers available (C′ is close to the vertical axis) they remain self-employed instead. Economies converging to C′ have a small fraction of poor relative to middle and upper class people. Wages are therefore high, but high wages in turn make it unlikely that any one will remain poor for very long. Again, high project returns insulate the agents in M and U from state L.

The same point is underscored by looking at some dynamic trajectories. X′ and Y′ are two points in the lower triangle which are close to each other but have slightly different mixes of the different classes. Now, starting at X′, which has the relatively larger middle class, the trajectory (which is drawn to be very close to the trajectory in the unperturbed model) moves the economy in a direction in which the ratio of the upper class to the lower class increases over time until the wages start rising. At the higher wages, however, workers can accumulate enough to become self-employed and therefore the lowest class declines very sharply and the economy converges to a situation with almost everybody self-employed.

The trajectory which starts at Y′ also moves in the same direction initially, but since the initial fraction of laborers was relatively large, wages do not rise and employing people remains profitable. Instead, the economy ends up at F′, which is a situation with both self-employment and entrepreneurial production.

If we identify self-employment with self-sufficient peasants and cottage industries and entrepreneurial production with large-scale capitalist agriculture and factory production, the dynamic patterns we describe above have interesting historical parallels. The most famous of these might be the instance of England and France, which in terms of the level of development and technology were roughly comparable at the middle of the 18th century (Crafts, 1977) and yet went through radically different paths of development. England went on to develop and benefit hugely from the factory system and large-scale production,while France remained a nation of small farms and cottage industries for the next hundred years. In terms of our model, one possible explanation would be that England started at a point like Y′ while France started at a point like X′.

While a full examination of the relevant historical data would be the subject of another paper, there seems to be at least some evidence that land was less equally distributed in England (where the enclosure movement had generated a large fraction of landless poor) than in France, particularly after the Revolution. Soltow (1968), Grantham (1975) and Clapham (1936) provide data on land ownership (albeit for later in the nineteenth century, although the distribution had not changed significantly over the preceding hundred years) which does suggest somewhat greater equality in France than in England. Clapham's discussion also indicates that the wage workers in France were indeed those who owned little or no land. Moreover, also consistent with our model were "the widespread complaints of the larger proprietors, that the existence of peasant property led to idleness, and prevented them from getting all the labour that their estates required" (p. 18). Brenner (1976) goes further to argue that the predominance of agricultural wage work in England, as opposed to peasant self-employment in France, led to the former's earlier development, although his evidence is somewhat sparse. It may be possible, then, to provide at least a partial explanation of this famous divergence using our model.

4. Conclusions

We have investigated an extremely rudimentary model of how the development of an economy into one or another institutional form can depend on initial conditions, specifically on the initial distribution of wealth. We conclude with a brief suggestion of directions for future research.

First, of course, is an inquiry into the mathematical structure of the problem. It is rare that models such as the one we have investigated should lend themselves to the kind of dimensional reduction we have engaged in.Nonetheless, models in which the state space is properly taken to be a set of distributions over individual characteristics, rather than simply the set of characteristics themselves, would seem to be easily generated in many applications, and it is already apparent that they may yield many interesting kinds of behavior.

One aspect of their behavior which we have not emphasized is that they provide a way to capture the empirically appealing notion that all possible individual characteristics (e.g. wealth levels) can be observed at any given time under any of several stationary distributions, (each distribution giving rise to different macro characteristics such as average wealth or fraction of the population in an occupation). For instance, it is probably reasonable to assert that for all practical purposes, the very richest people in India are almost as wealthy as the very richest in the United States; and the very poorest in the U.S. are no wealthier than their Indian counterparts. Yet standard Markov processes and deterministic representative agent models which give rise to multiple equilibria or hysteresis preclude this possibility: any state observed under one stationary distribution cannot be observed under another, so that if India and the U.S. correspond to different equilibria of the same standard model, then no Indian can enjoy the same wealth as any American.[14]

Second, we might consider how the various capital market imperfections which yield differential returns to wealth hold up as the economy grows and most agents become wealthier. Certainly, the story will be largely unchanged

---

[14]This is distinct from the idea, already common in economics (examples are Loury, 1981; Banerjee and Newman, 1989; Hopenhayn, 1989; Durlauf, 1990) that a stationary economy is one in which aggregate characteristics are fixed, but individuals may occupy different states over time.

if the minimum efficient scale rises with average wealth: then self-employment, while not out of the reach of anyone, will simply not be desirable because it won't yield a large enough return compared to the high wages available in larger enterprises.

Third, we could investigate the relationship between what we have called the institutional structure and growth through technological change. If, as suggested in the Introduction, the factory is more conducive to the development of new technology than the cottage (see North, 1981 for an elaboration of this view), then the dependence of the long run occupational structure on the initial distribution of wealth will also have implications for the growth path which the economy follows. For instance, suppose, as in much of the recent growth literature, that technological change is brought about largely through external or spill over effects, but that these effects differ under different institutional forms. Then individuals' occupational choice decisions will go largely as we have modeled them, and we might then find that the initial distribution of wealth can have consequences not only for the institutional structure, but also for the whole history of growth and technological change in the economy.

REFERENCES

Andreoni, J. (1989), "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, vol. 97, 1447-1458.

Banerjee, A.V. and A.F. Newman (1989), "Risk-bearing and the Theory of Income Distribution," CMSEMS working paper no. 877, Northwestern University.

Bernanke, B. and M. Gertler (1989), "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review* vol. 79, no. 1, 14-31.

——— (1990), "Financial Fragility and Economic Performance," *Quarterly Journal of Economics*, vol. CV, no. 1, 87-114.

Brenner, R. (1976), "Agrarian Class Structure and Economic Development in Pre-Industrial Europe," *Past and Present*, no. 70, 30-74.

Clapham, J.H. (1936), *Economic Development of France and Germany, 1815-1914*, 4th ed., Cambridge: Cambridge University Press.

Cohen, J.S. (1981), "Managers and Machinery: An Analysis of the Rise of Factory Production," *Australian Economic Papers*, vol. XX, no. 36, 24-41.

Crafts, N.F.R. (1977), "Industrial Revolution in England and France: Some Thoughts on the Question 'Why Was England First?'," reprinted in J. Mokyr, ed., *The Economics of the Industrial Revolution*, Totowa, New Jersey: Rowman & Allanheld, 1985.

Durlauf, S.N. (1990), "Locally Interacting Systems, Coordination Failure, and the Behavior of Aggregate Activity," mimeo Stanford University.

Eswaran, M. and A. Kotwal (1989), "Why are Capitalists the Bosses?," *Economic Journal*, vol. 99, no. 94, 162-76.

Galor, O. and J. Zeira (1988), "Income Distribution and Investment in Human Capital: Macroeconomic Implications," mimeo.

Grantham, G.W. (1975), "Scale and Organization in French Farming, 1840-1880," in W.N. Parker and E.L. Jones, eds. *European Peasants and Their Markets*, Princeton: Princeton University Press.

Hopenhayn, H.A. (1989) "A Dynamic Stochastic Model of Entry and Exit to an Industry," mimeo, Stanford University

Kehoe, T.J. and D.K. Levine (!989), "Debt Constrained Asset Markets," mimeo, Federal Reserve Bank of Minneapolis.

Khan, M.A. (1989), "In Praise of Development Economics," working paper no. 237, Johns Hopkins University.

Kihlstrom, R.E. and J.-J. Laffont, "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion," *Journal of Political Economy*, vol. 87, 719-748.

Loury, G.C. (1981), "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, vol. 49, 843-867.

Lucas, R.E. (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics*, vol. 22, 3-42.

MacDonald, G.M. (1982), "Specific Investments and Nonlabor Income," *Bell Journal of Economics*, vol. 13, no. 1, 225-233.

Matsuyama, K. (1990), "Increasing Returns, Industrialization and Indeterminacy of Equilibrium," CMSEMS working paper no. 878, Northwestern University.

Millward, R. (1981), "The Emergence of Wage Labor in Early Modern England," *Explorations in Economic History*, vol. XVIII, no. 1, 21-39.

Murphy, K.M., A. Shleifer and R.W. Vishny (1989a), "Income Distribution, Market Size, and Industrialization," *Quarterly Journal of Economics*, vol. CIV, no. 3, 537-564.

——— (1989b), "Industrialization and the Big Push," *Journal of Political Economy*, vol. 97, 1003-1026.

Newman, A.F. (1990), "The Capital Market, Inequality and the Prevalence of the Employment Relation," mimeo, Northwestern University.

North, D.C. (1981), *Structure and Change in Economic History*, New York: Norton.

Romer, P.E. (1986), "Increasing Returns and Long-run Growth," *Journal of Political Economy*, vol. 92, 1002-1037.

Sappington, D. (1983), "Limited Liability Contracts between Principal and Agent," *Journal of Economic Theory*, vol. 29, 1-21.

Soltow, L. (1968), "Long-run Changes in British Income Inequality," reprinted in A.B. Atkinson, ed., *Wealth, Income and Inequality*, Oxford: Oxford University Press, 1980.

Stiglitz, J.E. (1988), "Economic Organization, Information, and Development," in H. Chenery and T.N. Srinivasan, eds., *Handbook of Development Economics*, vol. I, fch. 5, New York: North-Holland, 93-160.

Stokey, N.L. and R.E. Lucas (1989), *Recursive Methods in Economic Dynamics*, Cambridge, Massachusetts: Harvard University Press.

Townsend, R. (1987), "Models as Economies," *Economic Journal*, vol. 98, no. 390, 1-21.