NBER WORKING PAPERS SERIES

WAS THE GREAT DEPRESSION A LOW-LEVEL EQUILIBRIUM?

John Dagsvik

Boyan Jovanovic

Working Paper No. 3726

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1991

We thank the National Science Foundation and the C.V. Starr Center for Applied Economics at New York University for technical and financial help. Some of the work was done while Dagsvik visited the C.V. Starr Center in October of 1987. The model was estimated by Ulf Uttersrud of the Central Bureau of Statistics in Norway. We owe him much. We thank Ray Atje for capable research assistance. We thank members of the "Coordination Failures" group that commented on an earlier version of this paper at the NBER Summer Institute in 1988, as well those that took part in the many seminars at which this paper was given. Finally, we thank an anonymous referee of this journal for comments that greatly improved the paper's content. This paper is part of NBER's research program in Economic Fluctuations. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3726
June 1991

WAS THE GREAT DEPRESSION A LOW-LEVEL EQUILIBRIUM?

ABSTRACT

Was the Great Depression the outcome of a massive coordination failure? Or was it a unique equilibrium response to adverse shocks? More generally, do aggregates fluctuate partly because agents occasionally settle on inferior, low-level equilibria? These questions lie at the heart of the current disagreement over how one should view business cycles. This paper estimates an employment model with monetary and real shocks. In one region of the parameter-space the model yields uniqueness, while in the other it yields up to three equilibria. When more than one equilibrium exists, a selection rule is needed. The equilibrium selection rule that we use has a Markovian structure, but the money supply is denied a coordination role -- it can not affect the choice of the equilibrium point. The global maximum likelihood estimates lie in the uniqueness region, implying that instead of being a low-level, coordination-failure equilibrium, the Depression era was caused by movements in fundamentals only. This result held for each of the three subperiods (since 1900) for which the estimation was done, but the estimates are imprecise and the conclusions that we draw from them are tentative. The paper also computes the <u>local</u> maxima in the region of multiplicity, and here some of our estimates indicate that the years 1932 and 1933 would have exhibited low level equilibria had more than one equilibrium existed.

John Dagsvik
Statistisk Sentralburo
Box 8131 dep
N-Oslo 1
Norway

Boyan Jovanovic
Department of Economics
New York University
269 Mercer Street,
7th Floor
New York, NY 10003
and
NBER

Was the Great Depression the outcome of a massive coordination failure? Or was it a unique equilibrium response to adverse shocks? More generally, do aggregates fluctuate partly because agents occasionally settle on inferior, low-level equilibria? These questions lie at the heart of the current disagreement over how one should view business cycles. This paper proposes a methodology within which such questions can be addressed.

In the real business cycle model (Prescott 1986) aggregates fluctuate because unique equilibrium points are displaced by shocks to technology. These shocks are intrinsic because they affect production possibilities. When money also affects trading possibilities (Lucas 1972) then a change in the supply of money too is an intrinsic shock that causes aggregate fluctuations.[1]

The real and the monetary approaches to fluctuations have both met with skepticism. Many reject the real business-cycle assumption that reductions in output are caused by technological regress (Greenwald and Stiglitz , 1988). And, money does not seem to lead the cycle (Kydland and Prescott, 1990), or explain much of the output variation (Sims 1980).[2] The emerging literature on coordination failures, summarized recently by Howitt (1990), takes a different view. It seeks to explain at least some aggregate fluctuations by means of extrinsic shocks, shocks that act to select one of several equilibrium outcomes. While the equilibrium point need not be chosen by an extrinsic random device -- the choice could reflect the realization of some observed or unobserved intrinsic shock -- models with multiple equilibria can clearly generate more volatility because aside from any intrinsic shocks that they may contain, they can also

---

[1] In Lucas's model equilibrium is unique within the class of equilibria that he considers, although there may be others that he does not consider that are more complicated functions of the history.

[2] And the variation in output that money does cause often in the wrong direction (Todd, 1990).

include a random mechanism that causes jumps among equilibria.

Some analyses of aggregates allow for random jumps among regimes (e.g, Hamilton 1989), but few build an equilibrium model of the economy that generates regimes endogenously as branches of its equilibrium correspondence.[3] This is what we shall do here. We shall ask whether aggregate labor market fluctuations reflect, in part, jumps among three Pareto-ranked equilibria. The jumps are driven by an extrinsic Markov process. The model has two real shocks and a monetary shock. For some parameter values equilibrium is unique, while for others there are three equilibria. Surprisingly, the maximum likelihood estimates lie in the uniqueness region. This is surprising because the multiple equilibrium subset of the parameter space has more parameters: besides the structural parameters, there also are the parameters of the equilibrium choice mechanism. The selection parameters are estimated jointly with the structural parameters, in the spirit of Kiefer and Wolfowitz (1956), and they were estimated separately for three different periods. In each case they were located in the uniqueness region. Even the 1920-1940 period, chosen especially because it contained the Great Depression -- a period that many think was a separate regime, a breed apart from the rest -- yielded the same result.[4]

Our aim here is to answer a specific question. In proceeding, we shall ignore some potentially useful information (such as data on wages) that, for technical reasons, we could not deal with. We shall ignore the possibility that the money supply was endogenous. All this limits the value of this paper as an

---

[3] Exceptions are Butler and Michell (1989), and Manning (1989), although they proceed quite differently from us.

[4] The Great Depression was anomalous not only in its severity but also in its character. For instance, Table 12 of Greenwald and Stiglitz (1988) shows that in Germany, Australia, the Netherlands, the U.K. and the U.S., real wages were higher at the trough of the depression in 1932 than had been at its outset in 1929.

investigation into the causes of the Great Depression. The paper's contribution is therefore largely methodological: Questions such as the one we ask are likely to come up again and again, and the way that one goes about answering them will probably have to resemble what we do here.

1. Detecting Multiplicity.

What dynamic behavior, in general, would one look for to see if the data are driven by jumps among distinct equilibria? This question can be answered only with a specific model, and even there we now have examples in which one can not identify the presence of extrinsic uncertainty (Hamilton and Whiteman, 1985). One must therefore restrict the way that intrinsic and extrinsic data affect observables so as to get identification in the sense of eq. (2) of Jovanovic (1989) -- without this, even an infinite sample will not contain the answer we seek.

Deterministic models with multiple solutions suggest that data should cluster around distinct equilibrium points.[5] Methods for finding the number of distinct clusters have been extensively discussed in the statistical literature (Fukunaga 1972). Here we shall add structure to the problem, structure of the kind commonly found in economic models. Let y be an endogenous observable, x an exogenous observable, and ϵ an exogenous unobservable. The pair (x, ϵ) is then an aggregate shock. Modifying Cooper's (1987) exposition slightly, let $U(y, y', x, \epsilon)$ be the representative agent's payoff function in the state (x, ϵ) when he takes a decision y, and when all others take the decision y'. Let U be continuously differentiable and concave in y, and let $f(y', x, \epsilon)$ be the optimal action for the agent at hand. Assuming that the maximum is interior, we must have

$$(1) \qquad \frac{\partial U}{\partial y}(y, y', x, \epsilon) = 0 \qquad \text{at } y = f(y', x, \epsilon) .$$

[5] Unless, of course, there is a continuum of equilibria.

The set of symmetric Nash equilibria for the game indexed by (x, ϵ) is

(2) $$\psi(x, \epsilon) = \{ y \mid y = f(y, x, \epsilon) \} .$$

To say something about the observable implications of the restriction in equation (2) we shall now get more specific and assume that[6]

$$f(y, x, \epsilon) = x + \epsilon + (\frac{7}{16})y + (\frac{3}{2})y^2 - y^3 .$$

Figure 1 shows the nature of equilibrium at two distinct values of $x + \epsilon$.

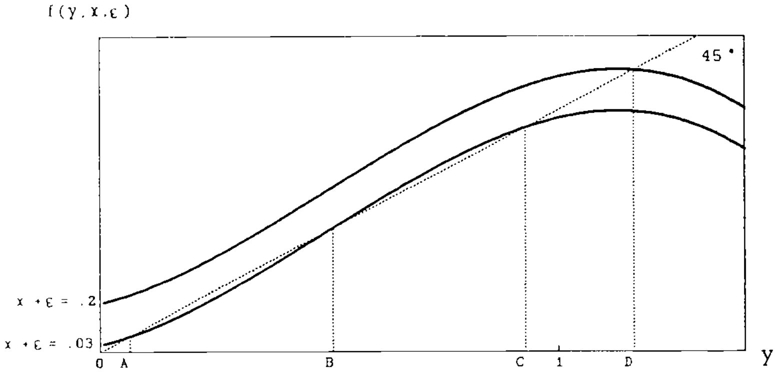

Figure 1: The determination of equilibrium

[6] It is hard to derive any observable implications in a general model with multiple equilibria. This was shown by Jovanovic (1989). One must make strong assumptions about the selection rule in particular, as Goldfeld and Quandt (1972) and Quandt (1972) pointed out some time ago. Also, see Kiefer (1978), Lee and Porter (1984), Coslett and Lee (1985) and Hamilton (1989). Lemma 1 of Jovanovic (1987) shows that for any f, a payoff function can be found that generates it as the optimal reaction function. The payoff function $U(y, y', x, \epsilon) = yf(y', x, \epsilon) - y^2/2$ is one (among many) that does the job. Obviously, any monotone transform of U would work too.

For some (x, ϵ) pairs there will be three equilibria, such as A, B, C, and for others there will be just one, such as D. The term $x+\epsilon$ shifts f up by the same amount for all y.[7] With these assumptions about f,

$$\psi(x,\epsilon) = \{ y \in R \mid x + \epsilon - (\frac{9}{16})y + (\frac{3}{2})y^2 - y^3 = 0 \} . \tag{3}$$

The equilibrium correspondence is therefore the set of solutions to a cubic equation in y. It is drawn in Figure 2.

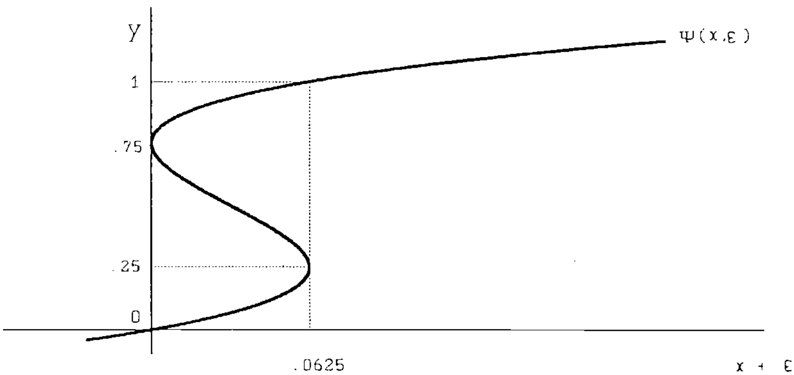

Figure 2: The symmetric equilibrium correspondence.

For $x+\epsilon \in (0, .0625)$ there are three equilibria. At $x+\epsilon = 0$ and $x+\epsilon = .0625$

[7] Many models have exactly three equilibria such as A, B, and C. This example will motivate the Phillips curve model that we shall estimate, and it will bring out some observable implications of multiplicity that should apply more generally.

there are exactly two, and everywhere else, equilibrium is unique. Now let's assume an equilibrium selection device that assigns equal probability to each branch of the equilibrium correspondence, conditional, of course, on $x+\epsilon$. So, for instance, when $x+\epsilon \in (0, .0625)$, each equilibrium branch gets a probability of 1/3. Let us also take ϵ to be distributed normally with mean zero and variance σ^2.

One way to summarize this model's implications is to look at the expectation of y conditional on x, $E(y|x)$, and its conditional variance $V(y|x)$. In doing so, we shall ask how fast information is lost as σ grows. The upcoming figures are based on a simulation that, for each value of x, drew 500 ϵ's and computed $E(y|x)$ and $V(y|x)$. Those ϵ's for which $x+\epsilon \in [0, .0625]$ also necessitated an equiprobable assignment to the 3 branches.

For $\sigma = .001$ the real side of the model is essentially deterministic, and all randomness stems from the selection device. That is why in Figure 3, $V(y|x)$ is positive only for $x \in [0, .0625]$, and why $E(y|x)$ jumps as x crosses the boundary of the region of multiplicity.

Figure 4 shows that these features of the data should remain visible as long as σ is an order of magnitude less than the range of the region of multiplicity. However, Figure 5 shows that as σ gets to be of roughly the same order as the magnitude of the region of multiplicity, the discontinuous-like behavior of both $E(y|x)$ and $V(y|x)$ disappears, although one can still see tremendous heteroskedasticity in y as x varies. Finally, as we increase σ even more, $E(y|x)$ becomes quite smooth, but the heteroskedasticity remains. This is shown in Figure 6. This example indicates that $V(y|x)$ is a better indicator of

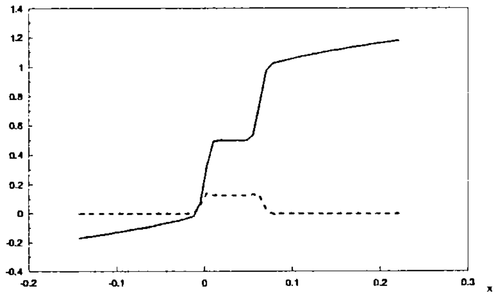

Figure 3: $\sigma = .001$.

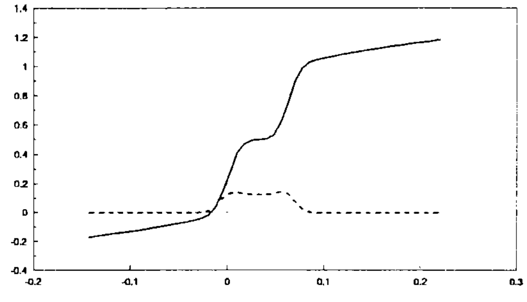

Figure 4: $\sigma = .01$

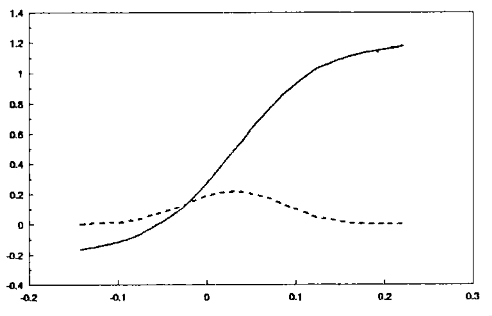

Figure 5: $\sigma = .05$

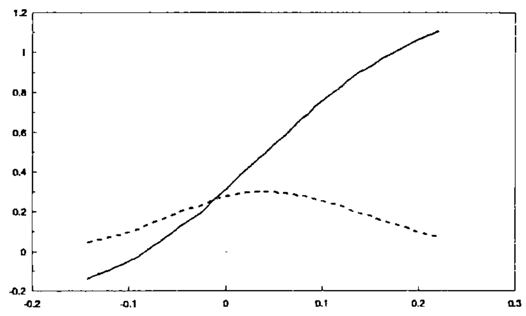

Figure 6: $\sigma = .1$

multiplicity than $E(y|x)$ if we expect a fair amount of noise. Indeed, this is what one would expect generally: If multiple equilibria arise in some region of x-space then switching among equilibria will create additional variance in y when x enters the region in question.

Why does this class of models exhibit multiplicity for intermediate values of $x+\epsilon$, and not at the extremely high and low values of $x+\epsilon$? The reason is that for tail values of $x+\epsilon$, playing high or playing low is a dominant strategy for a fraction of the population that is large enough to rule out multiplicity.

2. <u>A Macroeconomic Model of Employment</u>.

We shall estimate a model of employment with monetary shocks as well as real demand and supply shocks that has three equilibria for some parameter values. The only exogenous observable will be the money supply, or rather its surprise.

Agents differ in their fixed costs of employment c. Agent c takes action $a_c \in \{0,1\}$. If $a_c = 1$ he works, while if $a_c = 0$ he stays home. The expected payoff to working is $\phi(\bar{a},x,u)$, where x is the monetary surprise, u a shock to aggregate demand or productivity, and $\bar{a}$ the fraction of agents that work. The net payoff to agent c is

$$\begin{cases} \phi(\bar{a},x,u) - c & \text{if } a_c = 1, \text{ and} \\ 0 & \text{if } a_c = 0. \end{cases}$$

Being of measure zero, each agent takes $\bar{a}$ as given. And if his c is less than $\phi(\bar{a},x,u)$, he will play $a = 1$. Let v be an aggregate shock to the supply of labor and let $F(\cdot;\ v)$ be the distribution of c. Then a Nash equilibrium is an employment rate y for which

$$y = F[\phi(y,x,u);v]. \tag{3}$$

The model is static.[8] Its dynamics stem entirely from shocks to the supply of

---

[8] And it is much like Diamond's (1982) coconut example. Let coconuts grow on trees that cost c utils to climb, let F be the distribution of costs over trees, and let each person see one tree per period. Eating a coconut yields one util, while hunger yields zero utils. Eating one's own coconut is taboo, and the probability of meeting a trading partner is ϕ. The number of people climbing trees is "employment", and equilibrium employment solves (6).

money, and from shocks to the demand and the supply of labor.[9]

We now get specific about F and ϕ. Let $F(c;v) = 1/(1 + vc^{-\beta})$ for $c \geq 0$, where $\beta > 0$. An increase in v shifts the distribution of costs to the right, and therefore dampens the supply of labor.

We shall parametrize ϕ in such a way that two regimes will arise. In one regime, "regime U", equilibrium will be unique for all values of the exogenous data. In the other, "regime M", there will be three equilibria, for at least some values of the exogenous data. Equilibrium employment, y, is at the intersections with the 45° line in Figure 7, and for some parameter values there will be three such intersections.[10]

If ϕ were defined to be the real wage, this model would find it hard to ascribe low-level equilibrium status to the Great Depression, because the early years of the depression saw a rise in the real wage. Since equilibria are on the 45^0 line in Figure 7, and since F is increasing, this implies that employment and

[9] An infinite number of (ϕ, F) pairs can produce the same composition $F(\phi)$. Roger Klein has pointed out to us that this identification problem also shows up in discrete choice models; see Manski (1987). The problem can be solved if one has information on ϕ or F separately. For instance if one assumed that ϕ is the marginal product of labor and hence under price-taking also equal to the wage, then one could use wage-data to identify ϕ and F separately. But ϕ could incorporate queuing considerations, or synergies between market and non-market returns, or even lifetime returns that current wages do not capture. This, as well as the added complexity that a second endogenous variable would bring to the empirical work, led us to confine our analysis to just one endogenous variable (employment). On the other hand, Cogan (1981) estimates the mean of F at around 1,000 1966 dollars, although his estimate is for direct costs only, and excludes utility gained from nonmarket activities.

[10] For regime M to arise one must have a region in which ϕ increases rapidly enough with y. This could be because of trading externalities (Diamond 1982, Hall 1989), pecuniary externalities (Shleifer 1986, Murphy, Shleifer, and Vishny 1989), or technological externalities (Lucas 1988, Romer 1987). Increasing returns could also arise even in the short run if unused capacity varies with employment, as would happen if the short-run elasticity of substitution between capital and labor were close to zero. Papers that find increasing returns at the aggregate level at high frequencies include Caballero and Lyons (1990), Klenow (1990), and Walters (1963). To ensure that the curves in figure 7 start above the 45° line and end up below it, we need $\phi(0,x,u) > 0$ and $\phi(1,x,u) < \infty$ for all x and u.

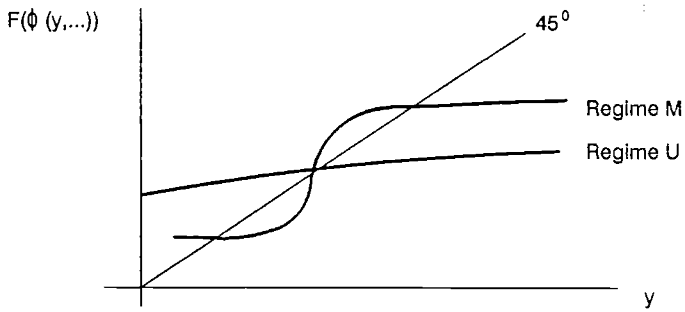

Figure 7: Determination of the equilibrium employment rate.

wages would have to be positively related.[11] In other words, equating ϕ to the real wage would make the Great Depression as big an outlier here as it is in models with unique equilibria.

Let $z = \log[y/(1-y)]$, so that[12] $y = 1/(1+e^{-z})$. Substituting for F and y in eq. (3) and inverting both sides yields

$$(4) \qquad e^{-z} = v\phi[(1+e^{-z})^{-1}, x, u]^{-\beta},$$

an implicit function to be solved for the equilibrium z. Now let $\phi(y, x, u) =$

[11] Which in developed countries they are, according to Table 7 of Greenwald and Stiglitz (1988).

[12] Note that $dz/dy = 1/y(1-y)$, so that z is a monotone increasing transform of y. In the estimation, we take y_t to be the employment/population ratio. Having obtained z in this way, we then linearly de-trend it. Employment, not hours, seems the appropriate measure for y since we vary only the extensive margin in this model. Kydland and Prescott (1989) have recently modelled employment variations in both the extensive and intensive margins. As for x_t, it is the surprise in the logarithm of the U.S. money-supply (M1) in period t, when it is regressed on its own lag, and lagged z. Exogeneity of x_t is, of course, doubtful, especially at yearly frequencies; we shall nevertheless assume it here.

$\hat{\phi}(y)x^{\delta}u$. Substituting for ϕ on the right-hand side of (4), taking logs and rearranging, we get

$$z - \beta\log\hat{\phi}[(1+e^{-z})^{-1}] = \beta\delta\log x + \beta\log u - \log v.$$

Now suppose that $\hat{\phi}[(1+e^{-z})^{-1}]$ well approximates a cubic in z.[13] Then

$$(5) \qquad z - \beta\log\hat{\phi}(\cdot) - E[\beta\log u - \log v] \approx \sum_{j=0}^{3}\alpha_j z^j .$$

Now if the two curves in figure 7 are to start above the 45^0 line and are to end up below it, $z - \beta\log\hat{\phi}$ must go to infinity as $z \rightarrow +\infty$, and to minus infinity as $z \rightarrow -\infty$.[14] Since z^3 is the dominant term on the right-hand side of (5), this means that α_3 must be positive.

Substituting in (5) and dividing by α_3 leads to:

$$(6) \qquad z^3_t + a_1 z^2_t + a_2 z_t + c\log x_t + a_0 = \epsilon_t$$

---

[13] The restriction that this places on $\hat{\phi}(y)$ is, of course, that $\hat{\phi}(y)$ approximate a cubic in the variable $\log[y/(1-y)]$, so that ϕ must be translog in form.

We chose the cubic approximation for several reasons. First, it delivers the familiar S-shape under regime M in Figure 7. Second, a quadratic leads to the problem of nonexistence of equilibrium for some values of x, u and v. And third, a polynomial of order higher than 3 would need to be of an order of at least 5 (for reasons having to do with corner conditions, reasons that we shall elaborate on below), and with our short time series that number of parameters could not be precisely estimated.

[14] Loosely speaking, if ϕ is some multiple of the marginal product of labor, this marginal product must go to infinity as $y \rightarrow 0$, and it must go to zero as $y \rightarrow 1$.

where $a_1 = \alpha_2/\alpha_3$, $a_2 = \alpha_1/\alpha_3$, $c = -\beta\delta/\alpha_3$, $a_0 = \alpha_0/\alpha_3$, and $\epsilon_t = ([\beta \log u - \log v] - E[\beta \log u - \log v])/\alpha_3$. Since β and α_3 are both positive, if we believe that $\delta > 0$ then at any locally stable equilibrium, we expect to find that $c < 0$. This point will resurface when we look at the empirical results.

Let us pause here and go over what we have done. With the specific functional forms, the transformation from y to z, and the cubic approximation for $\hat{\phi}$, we have transformed our structural equation (3) into the manageable form (6). The ideal procedure is to estimate this implicit function directly rather than solve it for the reduced form correspondence, because one would not then have to assume anything about how the solutions to (6) are chosen. Unfortunately the maximum likelihood method will not do for this purpose because when (6) has multiple roots, the value of the Jacobian of the transformation from ϵ to z will generally vary over the solutions to (6). So we used an instrumental variable method instead, but as the results were rather poor we shall discuss them only at the end.

We now specify the laws of motion for the unobservable variables -- ϵ and the equilibrium indicator. A large ϵ reflects a rightward shift of demand, or a rightward shift of supply, or both. As in the real business cycle model (Prescott 1986) let

(7) $$\epsilon_t = \rho\epsilon_{t-1} + \eta_t,$$

with η_t serially uncorrelated and distributed normally with mean zero and variance σ^2_η. Let

(8) $\lambda = a^2_1/9 - a_2/3$.

When $\lambda > 0$, let $d_1 = -2\surd\lambda - a_1/3$, and let $d_i = d_1 + i\surd\lambda$ for $i = 2, 3, 4$. Also, let $r_1 = -2\lambda\surd\lambda + a_0 - a_1a_2/3 + 2a^3_1/27$, and $r_2 = r_1 + 4\lambda\surd\lambda$.

The d_i and r_i are shown in figure 8. Multiple equilibria arise only when ϵ - ϵclogx $\in [r_1, r_2]$. As $\lambda \to 0$, $r_2 - r_1 \to 0$, and multiple equilibria disappear. In other words, we are in regime M if and only if $\lambda > 0$.

When $\lambda < 0$, the equilibrium correspondence is still S-shaped, but it is now single-valued.

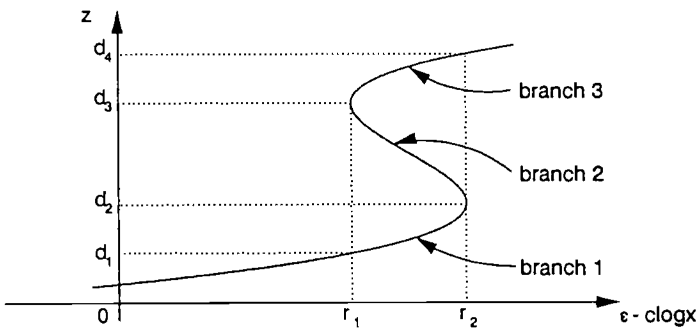

Figure 8: The equilibrium correspondence in regime M.

Autocorrelated equilibrium selection. When in regime M, and when ϵ - clogx (r_1, r_2), one of three equilibria must be chosen. In this region the equilibrium correspondence has three branches, which we label 1, 2, and 3 from the bottom up. While ϵ - clogx is in the region (r_1, r_2), we shall assume that selection obeys a 3 X 3 Markov chain. This Markov chain describes the link between equilibrium choices in consecutive periods.

But what if ϵ - clogx was not in (r_1, r_2) in the previous period, and enters it this period? Here, we assume that the probability that a branch is chosen is the steady state probability for that branch for the Markov chain described above. This is natural if one believes that the selection process has memory during consecutive periods in (r_1, r_2), but not if the spell in (r_1, r_2) is interrupted. Moreover, this assumption introduces no new parameters. All in all, the selection mechanism introduces three new parameters.

The selection mechanism can capture a variety of plausible hypotheses about how equilibria are chosen. For instance, if $\pi_2 = 0$, the "naively" unstable branch 2 will never materialize.[15] And if $\pi_3 = 1$, the economy will always settle on the high activity equilibrium. Under some added assumptions, this equilibrium is the one at which welfare is also at its highest, although further Pareto improvements could occur if employment were raised even further so long as this could be done without distorting a margin other than the one bearing on the participation decision.

Let $d = (d_1, d_2, d_3, d_4)$. Then d determines what branch of the equilibrium correspondence a particular z_t lies on. Let $Y^i{}_j(z_{t-1}, z_t, d) = 1$ if z jumps from branch i at t-1, to branch j at t, and let it be zero otherwise. Similarly, let $Y_j(z_{t-1}, z_t, d) = 1$ if z jumps to branch j at t from somewhere outside the region of multiplicity at t-1, and let it be zero otherwise. Then we shall assume that

[15] While we would prefer not to even mention out of equilibrium dynamics, one can hypothesize something similar to the familiar tatonnement process in Walrasian equilibrium. The middle solution is unstable in the sense that if "by mistake" employment were, say, a bit larger than the equilibrium value, even more people would wish to enter the market, and if it were a bit smaller, even more would wish to withdraw.

$$\Pr\{Y_j^i(z_{t-1}, z_t, d) = 1\} = \begin{cases} b\pi_j & \text{if } i \neq j \text{, and} \\ 1-b+b\pi_i & \text{if } i = j \text{,} \end{cases} \tag{9}$$

where $\sum \pi_j = 1$. Thus when $i \neq j$, the transition probabilities are independent of i. The steady state probabilities of this 3-state Markov chain are just π_1, π_2 and π_3, which means that

$$\Pr\,(\,Y_j(z_{t-1},\ z_t,\ d) = 1\,) = \pi_j, \qquad j = 1,\ 2,\ 3. \tag{10}$$

The 3 new parameters are π_1, π_2 and the parameter b, which is an index of serial dependence in the selection mechanism. When $b = 0$ no transitions occur (perfect autocorrelation). When $b = 1$ the selection is i.i.d., with probabilities π_1, π_2 and π_3.[16] One can thus think of ρ and 1-b as parameters that compete in explaining the persistence that the z_t exhibit. The first measures the persistence stemming from the real side, the second measures the persistence arising from the selection side.

<u>The Likelihood Function</u>. Let $I_{\{i=j\}} = 1$ if $i = j$ and zero otherwise. Then the log likelihood of the sample is

[16] A coordination role for money could be captured by allowing the π_i and possibly b as well to depend on x_t, or on M1 itself, in some simple way.

$$\log L = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{t=2}^{T} Y_j^i(z_{t-1}, z_t, d)\log(b\pi_j + I_{(i=j)}(1-b))$$

$$(11) \quad + \sum_{j=1}^{3}\sum_{t=2}^{T} Y_j(z_{t-1}, z_t, d)\log\pi_j - (T-1)(\log\sigma_\eta + \log\sqrt{2\pi})$$

$$+ \sum_{t=2}^{T}\log|3z_t^2 + 2a_1z_t + a_2| - \left(\frac{1}{2\sigma_\eta^2}\right)\sum_{t=2}^{T}\Big[z_t^3 + a_1z_t^2 + a_2z_t + a_0(1-\rho)$$

$$+ c\log x_t - \rho(z_{t-1}^3 + a_1z_{t-1}^2 + a_2z_{t-1} + c\log x_{t-1})\Big]^2 .$$

The likelihood factors so that one can maximize it in three steps. The first step conditions on d and chooses ρ, c, σ_η and a_0 to maximize

$$(12) \quad -(T-1)\log\sigma_\eta - (1/2\sigma^2_\eta)\sum_{t=2}^{T}(z^3_t + a_1z^2_t + a_2z_t + a_0(1-\rho)$$
$$+ c\log x_t - \rho(z^3_{t-1} + a_1z^2_{t-1} + a_2z_{t-1} + c\log x_{t-1})).$$

The Jacobian $\sum_{t=2}^{T}\log|3z_t^2 + 2a_1z_t + a_2|$ is left out at this stage, and the next because once d (and hence a_1 and a_2) is given, this expression is a constant. The second step still conditions on d and chooses π_1, π_2, π_3 and b to maximize

$$(13) \quad \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{t=2}^{T} Y_j^i(z_{t-1}, z_t, d)\log(b\pi_j + I_{(i=j)}(1-b)) + \sum_{j=1}^{3}\sum_{t=2}^{T} Y_j(z_{t-1}, z_t, d)\log\pi_j.$$

The functional forms of (9) and (10) prevent us from calculating the ML estimate of the π_i explicitly. A standard approach would then use the first order conditions along with an iteration procedure. However, we use a less costly alternative based on an algorithm that maximizes (13) over a large finite set of parameter values. This set is a discretized approximation of the parameter space generated by π_1, π_2 and b. The third and final stage maximizes the likelihood over d. A more detailed description of the procedure is in the Appendix.

Identification. When $\lambda < 0$, this problem is standard and the selection rule in eq. (9) and eq. (10) does not come into play. But when a_1 and a_2 imply a positive λ, they determine the d_i and hence the Y_j^i and Y_j uniquely, and this identifies b and the π_j. The likelihood function is otherwise standard, and its remaining parameters clearly are identified.

Estimates. We computed maximum likelihood estimates using annual data on employment and the money-supply for three separate time periods: 1901-1940, 1921-1940 and 1951-1986. Surprisingly, uniqueness does better than multiplicity in each case -- L_{max} is highest when $\lambda < 0$. We also computed the local maximum, in each period, over the region for which $\lambda > 0$. These local maxima are reported in the first column of each table. A curious and unexpected property of the estimates for the general case is that when λ is in the region where the structure is afforded the "luxury" of assigning observations into one of three separate regimes, it chooses to assign all of the observations to the second, middle branch. This is all the more puzzling because this branch is unstable.

Surprise money does well in explaining the movements in employment. For instance, employment reductions during the Great Depression coincide with periods in which surprise money was negative. This seeming ability of the exogenous regressor to "explain" what would otherwise appear as shifts in regime might preclude the structures with multiplicity from doing better. For this reason, columns 4 and 5 of each table report estimates that constrain c to be zero.[17]

Following that, we constrain ρ to be zero. Under this constraint, any persistence that z_t exhibits is explainable only by persistence in the way equilibria are chosen. These estimates too should favor multiplicity. Finally, we report estimates that set ρ and c both equal to zero. This case should favor multiple equilibria the most.

Turning to table 1, we find that the estimates that the first column reports do not maximize the likelihood. Moreover, at the local maximum in the $\lambda > 0$ region, $\pi_2 = 1$ -- the selection rule always chooses the middle, unstable branch! Because of this, c is estimated to be positive, implying that δ is negative, so that surprise money reduces ϕ ! This rather strange estimate comes about because the raw correlation between z and x is positive (causing a highly significant and positive OLS coefficient reported at the foot of the table), and the only way to get the model to mimic this at an unstable equilibrium is if c is in fact positive. This is clear from figure 8. The estimate of $b = 1$ is meaningless: It says that equilibrium selection is iid, but this is because selection is driven by a degenerate random variable, so that b is not identified.

The ML estimate is in column 2. It lies in the uniqueness region, and the equilibrium is stable, which is reflected in a negative estimate of c. Figure 9a plots the estimated equilibrium correspondence at the median ϵ, as a function of x. This is the curve ML. The curve $\widetilde{ML}$ is based on the first column of the table. Figure 9a also depicts the linear OLS regression line. The figure is deceptive in one respect: At first glance it would seem that a cubic could do

[17] We present these estimates to help us understand the estimation techniques, and not because we expected the restrictions $c = 0$ to be valid; clearly, it isn't because it denies any relation between x and z, and yet the data show a significant and positive correlation between the two in each subperiod.

Table 1: Results for 1901-1940

	The General Case			The Case c = 0			The Case ρ = 0			The Case c=0 and ρ=0		
	$\lambda>0$	$\lambda<0$	t	$\lambda>0$	$\lambda<0$	t	$\lambda>0$	$\lambda<0$	t	$\lambda>0$	$\lambda<0$	t
MxL	109.58	112.81		101.30	104.74		93.19	95.79		92.45	91.72	
d_1	-0.721			-0.469			-0.262			-0.253		
d_4	0.770			-.469			0.336			0.325		
b	1			.01			0.31			0.36		
π_1	0			0			0.1			0.08		
π_2	1			0			0.9			0.92		
π_3	0			1			0			0		
ρ	0.756	0.757	7.1	0.691	0.70	5.7						
c	0.216	-0.05	-1.7				0.0102	-0.037	-1.8			
σ_η	0.016	0.0038	1.8	0.032	0.004	2.1	0.003	0.0047	2.2	0.0029	0.0051	2.3
a_0	0.0065	-0.0019	-0.7	-0.016	-0.0006	-0.3	0.0014	-0.0015	-1.5	0.0013	-0.0077	-1.6
a_1	-0.074	0.199	2.5	1.407	0.278	3.1	-0.111	0.277	3.1	-0.108	0.3204	3.5
a_2	-0.415	0.0823	1.7	0.660	0.072	1.8	-0.063	0.067	1.8	-0.059	0.066	2.0
$\tilde{\lambda}$	0.139	-0.023		0.0	-0.015		0.022	-0.014		0.021	-0.011	

Least Squares Estimates (t-ratios in parentheses):

$$z = .004 + .59 \log x \qquad R^2 = .26$$
$$\quad (.37) \qquad (3.66)$$

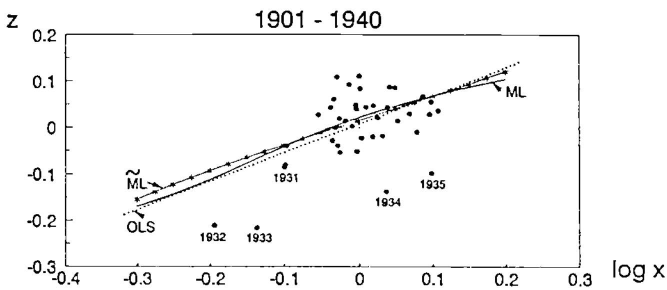

Figure 9a: Equilibrium at ϵ = 0

better than ML in that it could get closer to some of the points if it had more curvature. While this is true at $\epsilon = 0$, it would not maximize the likelihood because ϵ bounces around.[18]

The ML estimate remains in the uniqueness region even if we force ρ or c to be zero one at a time. Only when <u>both</u> of them are set to zero does the ML estimate choose the parameters from the region of multiplicity. A graphical representation in z - logx space is not helpful when $c = 0$ because z and x are then unrelated. But the case where $\rho = 0$ is of some interest: When $\rho = 0$, the selection rule now assigns positive weight to the bottom branch of the equilibrium correspondence, although most of the weight still goes on the middle branch. Just two out of the forty z_t's are put on the bottom branch of the equilibrium correspondence. The years were 1932 and 1933. The estimated equilibrium correspondence is plotted in figure 9b.

According to these estimates, the autocorrelation in the choice of equilibria, as measured by 1-b is high: The probability of remaining on branch i from one period to the next is $1\text{-}b\text{+}b\pi_i$. The last two sets of estimates in table 1 put this probability at about .97 for branch 2, and at about .7 for branch 1. We had originally conjectured that such parameter values would be part of the <u>unconstrained</u> ML estimate, but this was not the case.[19]

---

[18] The stationary distribution of ϵ has variance $\sigma^2_\eta/(1-\rho^2)$ which, according to the estimates in column 2, means that the standard deviation of ϵ is about .03. On the other hand, the estimates of σ_ϵ implied by the first column are three times as high. According to figure 8 this estimate must be divided by c to correct it into units of x. Using estimates of (ρ, σ_ϵ) from either column gives us a standard deviation of roughly unity for the effect of unobservables in units of x.

[19] Bernanke (1983, esp. Table 2) observes that his <u>linear</u> Phillips curve-type model leaves much of the depression-era variability in output unexplained. His response is to add variables measuring bank failures and liabilities of firms on the right hand side of his regressions, but the endogeneity of these variables makes those regressions hard to interpret. Our approach is to instead add nonlinearity to the underlying relationship between employment and surprise money.

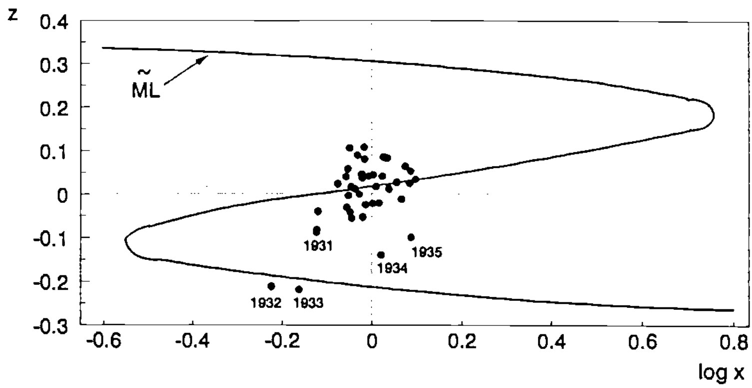

Figure 9b: 1901-1940 $\tilde{ML}$ estimate at $\epsilon = 0$, constrained by $\rho = 0$.

The next set of results deal with the 1921-40 period, and they are reported in table 2. We had expected structures with non-uniqueness to do especially well here. That they do not in the general case reflects surprise money's ability to track employment fairly closely in this subperiod. But as soon as either ρ or c are constrained at zero, multiple equilibrium is the ML choice. But the middle branch still gets all the weight ($\pi_2 = 1$). Paradoxically, this middle branch of the ML curve slopes in the opposite direction of the ML and the OLS line (see figure 10), which is no doubt caused by the large sampling error in the estimated coefficients. The estimate of b is meaningless here because there are no transitions among the three branches, and so b plays no role.[20] Figure 10 plots

[20] In eq. (13), when $\pi_2 = 1$, the expression $b\pi_j + I_{\{i=j\}}(1-b)$ does not depend on b when $j = 2$. When $j \neq 2$, Y_j^i is itself zero, so again the likelihood does not depend on b.

Table 2: Results for 1921-1940

	The General Case			The Case $c = 0$			The Case $\rho = 0$			The Case $c=0$ and $\rho=0$		
	$\lambda>0$	$\lambda<0$	t	$\lambda>0$	$\lambda<0$	t	$\lambda>0$	$\lambda<0$	t	$\lambda>0$	$\lambda<0$	t
MxL	55.002	56.155		53.732	51.923		54.65	49.39		53.373	46.238	
d_1	-0.199			-0.199			-0.198			-0.199		
d_4	0.056			0.056			0.0555			0.056		
b	0.91			0.91			0.91			0.91		
π_1	0			0			0			0		
π_2	1			1			1			1		
π_3	0			0			0			0		
ρ	0.194	0.750	4.8	0.223	0.713	4.1						
c	-0.0075	-0.034	-2.1				-0.0070	-0.023	-2.3			
σ_η	0.0012	0.0024	2.5	0.0013	0.0031	2.4	0.0012	0.0024	2.9	0.0013	0.0031	2.9
a_0	-0.0013	-0.001	-0.4	-0.0014	-0.0028	-0.9	-0.0013	-0.001	-1.2	-0.0013	-0.0014	-1.5
a_1	0.215	0.138	2.7	0.215	0.198	3.1	0.215	0.169	3.5	0.215	0.209	3.6
a_2	0.0032	0.328	1.6	0.0032	0.0364	1.7	0.0032	0.0196	1.7	0.0032	0.0253	1.9
λ	0.004	-0.009		0.004	-0.008		0.004	-0.003		0.004	-0.004	

Least Squares Estimates (t-ratios in parentheses)

$$z = \underset{(.55)}{.01} + \underset{(2.45)}{.59} \log x \qquad R^2 = .25$$

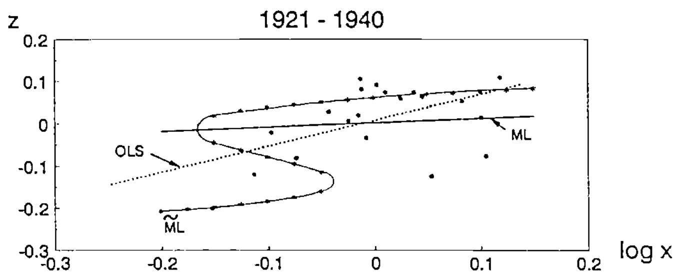

Figure 10: Equilibrium at $\epsilon = 0$

ML and $\tilde{ML}$ (for the general case) along with the linear OLS regression line.

Table 3 reports results for the 1951-86 period, results that we anticipated would not show evidence of multiplicity because this set of bivariate data seem to contain no outliers, no distinct clusters. As the table shows, uniqueness does better than multiplicity no matter what restrictions one imposes. Incidentally, the OLS regressions show that the Phillips curve relation was weaker in this period than prior to World War II.

On the whole, then, employment fluctuations since 1900 are better accounted for by a structure that exhibits uniqueness of equilibrium at all times, a structure that ascribes all fluctuations to movements in fundamentals. But this conclusion is tenuous because λ does not differ significantly from zero. Its ML estimates are -0.023, -0.009, and -0.023 for the three subperiods. Given the standard errors for a_1 and a_2, these estimates of λ do not differ significantly from zero.

The Instrumental Variables Approach. When estimating models that have multiple solutions, the instrumental variables approach is appealing because it requires neither an explicit solution for the endogenous variable (See Amemiya (1973)), nor any assumption about the way in which solutions are chosen. That is, the assumptions we make in equations (9) and (10) are unnecessary with this approach, which works directly on the implicit function defining the equilibrium point, which in this case is in eq. (6).

We used the approach as described in Judge et al (1985), p. 168, to estimate the reduced form equation (6) for the special case where $c = 0$:[21]

---

[21] A more advanced treatment of these issues is in Gallant (1977).

Table 3: Results for 1951-1986

	The General Case			The Case c = 0			The Case ρ = 0			The Case c=0 and ρ=0		
	$\lambda>0$	$\lambda<0$	t	$\lambda>0$	$\lambda<0$	t	$\lambda>0$	$\lambda<0$	t	$\lambda>0$	$\lambda<0$	t
MxL	79.424	81.364		79.400	81.332		75.824	78.605		74.106	76.845	
d_1	-0.578			-0.578			-0.640			-0.636		
d_4	0.589			0.588			0.604			0.600		
b	1			1			1			1		
π_1	0			0			0			0		
π_2	1			1			1			1		
π_3	0			0			0			0		
ρ	0.498	0.461	2.7	0.511	0.478	3.2						
c	0.04	-0.022	-0.3				0.404	-0.099	1.4			
σ_η	0.016	0.0058	1.6	0.016	0.0057	1.7	0.021	0.005	1.9	0.021	0.0052	1.9
a_0	0.001	-0.0002	-0.1	0.001	0.00001	0.01	0.00067	0.0002	0.2	0.00044	0.00026	0.3
a_1	-0.0165	-0.0453	-0.5	-0.015	-0.0718	-0.9	0.054	-0.0953	-1.3	0.0540	-0.0953	-1.3
a_2	-0.255	0.0684	1.4	-0.255	0.0672	1.4	-0.289	0.0526	1.5	-0.286	0.0526	1.5
λ	0.085	-0.023		0.085	-0.022		0.097	-0.017		0.095	-0.017	

Least Squares Estimates (t-ratios in parentheses)

$$z = \underset{(.4)}{.005} + \underset{(1.97)}{1.55} \log x \qquad R^2 = .10$$

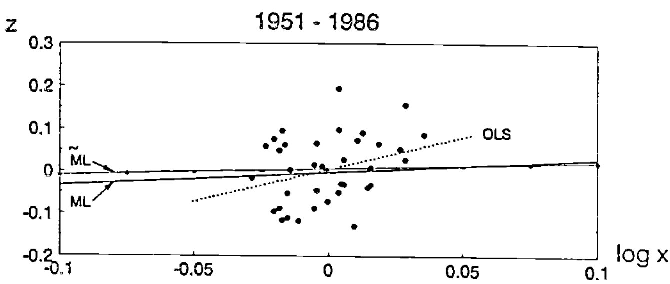

Figure 11: Equilibrium at $\epsilon = 0$

(6)' $$z_t^3 = -a_0 - a_2 z_t - a_1 z_t^2 - \epsilon_t .$$

We used x_t and x^2_t as instruments for z_t and z^2_t, respectively, in eq. (6)', while treating z^3_t as the dependent variable. Since we use x_t as one of the instruments, we can not include it as an additional regressor in (6)' since its coefficient, c , would not be identified.

The results are in Table 4. They show that the approach works poorly here. The estimates are quite imprecise; in most cases their absolute values are less than the standard error. The reason for this is of course that x_t and x^2_t are poor instruments for z_t and z^2_t, so that this procedure also offers no firm conclusion about the sign of λ.

Table 4: Instrumental Variables Estimates

	1901-1940		1921-1940		1951-1986	
	est.	std. error	est.	std. error	est.	std. error
a_0	-.00136	.00268	.00176	.00984	.00347	.0055
a_1	.30542	.43176	-.11132	1.0531	-.57139	.87505
a_2	.00999	.06397	-.04129	.1239	-.00446	.01474

3. Conclusion.

In spite of the extra parameters that the multiple equilibrium structures introduce, namely π_1, π_2, and b, they fit about the same or a little worse than structures with unique equilibria. The exercise does not, however, treat multiplicity in the most charitable way because the mechanism that chooses equilibria does not depend on x. An S-shaped curve passing through the points in figures 10, 11 and 12 will explain much more if x_t is allowed to determine which of three branches gets chosen at date t. This will be the case if the Fed's actions serve a coordination role and act to signal which equilibrium gets chosen. But how does one operationalize such a concept? If the selection rule's dependence on x_t is arbitrary, this introduces as many parameters as there are time periods, and with just two time-series on observables, the loss of degrees of freedom is too large. Moreover, section 5B of Jovanovic (1989) shows that unless one has prior information on just how exactly an exogenous observable affects the selection mechanism, one is, in this bivariate context at least, likely to fall prey to using a model that is consistent with (or, rather, can not be refuted by) just about any correlation pattern one might see. The natural procedure would use a parsimonious representation of the π_i and b as a function of x_t or x_{t-1} or other exogenous observables.

REFERENCES

Amemiya, T. "Regression Analysis When the Dependent Variable is Truncated Normal," Econometrica 41, no. 6 (November 1973): 997-1016.

Bernanke, B.S. "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression," American Economic Review 73, no. 3 (June 1983): 257-76.

Butler, J.S., and J.M. Mitchell, "A Simultaneous Equation Model of Cardiovascular Disease and Earnings," unpublished paper, Vanderbilt University, 1989.

Caballero, R. and R. Lyons, "Externalities and the Solow Residual," unpublished paper, Columbia University, August 1990.

Cogan, J.F. "Fixed Costs and Labor Supply," Econometrica 49, no. 4 (July 1981): 945-63.

Cooper, R. "Dynamic Behavior of Imperfectly Competitive Economies with Multiple Equilibria," NBER working paper 2388, September 1987.

Cosslett, S., and L-F. Lee, "Serial Correlation in Discrete Variable Models," Journal of Econometrics, 27(1985): 79-97.

Diamond, P "Aggregate Demand Management in Search Equilibrium," Journal of Political Economy 90, no. 5 (October 1982): 881-94.

Fisher, F. The Identification Problem in Econometrics, McGraw-Hill: New York 1966.

Fukunaga, K. Introduction to Statistical Pattern Recognition, New York: Academic Press, 1972.

Gallant, R.A. "Three-Stage Least-Squares Estimation for a System of Simultaneous, Nonlinear, Implicit Equations," Journal of Econometrics, 5 (1977): 71-88

Goldfeld, S. and R. Quandt Nonlinear Methods in Econometrics, North Holland: Amsterdam 1972.

Greenwald, B. and J. Stiglitz "Examining Alternative Macroeconomic Theories," Brookings Papers on Economic Activity 1 (1988): 207-60.

Hall. R. "Temporal Agglomeration," NBER working paper no. 3143, October 1989.

Hamilton, J.D. and C.H. Whiteman, "The Observable Implications of Self-Fulfilling Expectations," Journal of Monetary Economics, 16, no. 3 (November 1985): 353-74.

Hamilton, J.D. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," Econometrica, 57, no. 2 (March 1989): 357-84.

Howitt, P. "Determinate Outcomes with Multiple Equilibria," unpublished paper, University of Western Ontario, June 1990.

Jovanovic, B. "Micro Shocks and Aggregate Risk," Quarterly Journal of Economics, (May 1987): 395-409.

Jovanovic, B. "Observable Implications of Models with Multiple Equilibria," Econometrica, 57, no. 6 (November 1989): 1431-38.

Judge, G., W.E. Griffiths, R. Hill, H. Lutkepol and T-C Lee. The Theory and Practice of Econometrics, New York: John Wiley, 1985.

Kiefer, J. and J. Wolfowitz, "Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters," Annals of Mathematical Statistics, 27 (1956): 887-906.

Kiefer, N. "Discrete Parameter Variation: Efficient Estimation of a Switching Regression Model," Econometrica 46, no. 2 (March 1978): 427-34.

Klenow, P. "Externalities and Economic Fluctuations," unpublished paper, Stanford University, July 1990.

Kydland, F. and E.C. Prescott, "Hours and Employment Variations in Business Cycle Theory," Discussion Paper no. 17, Federal Reserve Bank of Minneapolis, August 1989.

Kydland, F. and E.C. Prescott, "Business Cycles: Real Facts and a Monetary Myth," Federal Reserve Bank of Minneapolis Quarterly Review, (Spring 1990): 3-18.

Lee, L-F, and R.H. Porter, "Switching Regression Models with Imperfect Sample Separation Information - With an Application to Cartel Stability," Econometrica, 52, no. 2 (March 1984): 391-418.

Lucas. R.E. Jr. "Expectations and the Neutrality of Money," Journal of Economic Theory, (1972).

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

Manning, A. "Multiple Equilibria in the British Labour Market," unpublished paper, London School of Economics, September 1989.

Manski, C. "Identification of Binary Response Models," University of Wisconsin, August 1987.

Murphy, K., A. Shleifer, and R. Vishny, "Industrialization and the Big Push," Journal of Political Economy, 97, no. 5 (October 1989): 1003-26.

Prescott, E.C. "Theory Ahead of Business-Cycle Measurement," Carnegie-Rochester Series on Public Policy, 25 (1986): 11-44.

Quandt, R.E. "A New Approach to Estimating Switching Regressions," Journal of the American Statistical Association, 67, no 338 (June 1972): 306-10.

Romer, P. "Crazy Explanations for the Productivity Slowdown," NBER Macroeconomic Annual, 1987.

Shleifer, A. "Implementation Cycles," Journal of Political Economy, 94, no. 6 (December 1986).

Sims, C. "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered," American Economic Review, 70 (May 1980): 250-57.

Todd, R. "Vector Autoregression Evidence on Monetarism: Another Look at the Robustness Debate," Federal Reserve Bank of Minneapolis Quarterly Review, (Spring 1990): 19-37.

Walters, A. "A Note on Economies of Scale," Review of Economics and Statistics, (November 1963): 425-27.

Appendix: The Maximum Likelihood Procedure. Eq. (7) implies that

$$z_t^3 + a_1 z_t^2 + a_2 z_t + a_0(1-\rho) + c\log x_t - \rho(z_{t-1}^3 + a_1 z_{t-1}^2 + a_2 z_{t-1} + c\log x_{t-1}) = \eta_t$$

where $\{\eta_t\}$ is white noise with η_t normally distributed $N(0, \sigma_\eta)$. When equilibrium is unique, i.e., when ϵ_t-$c\log x_t \notin [r_1, r_2]$, the likelihood therefore is of the form

$$L_1 = |J| \; (\sigma_\eta\sqrt{2\pi})^{-T+1} \exp[-\frac{1}{2\sigma_\eta^2}\sum_{t=2}^{T} (z_t^3 + a_1 z_t^2 + a_2 z_t + a_0(1-\rho) + c\log x_t$$
$$-\rho(z_{t-1}^3 + a_1 z_{t-1}^2 + a_2 z_{t-1} + c\log x_{t-1}))]$$

where J is the Jacobian of the mapping $(\epsilon_2, \epsilon_3, \ldots, \epsilon_T) \rightarrow (Z_2, Z_3, \ldots, Z_T)$, and where $Z_T = z_t^3 + a_1 z_t^2 + a_2 z_t + c\log x_t + a_0$. By straight forward calculus

$$\log|J| = \sum_{t=2}^{T} \log|3z_t^2 + 2a_1 z_t + a_2| \; .$$

In the regime with three equilibria we must take into account the likelihood of the selection of equilibria. Let q_{ij} be the probability of switching from branch i to branch j from one year to another, i, j = 1, 2, 3. Similarly let P_j be the probability of selecting branch j, j = 1, 2, 3, given a unique equilibrium in the previous year.

From assumptions (9) and (10), $q_{ij} = \beta\pi_j + (1-\beta)I_{\{i=j\}}$ and $P_i = \pi_j$. It is easily verified that $\{\pi_j\}$ are the steady state probabilities of this Markov chain.

The Likelihood of the equilibrium selection process, L_2, is therefore

$$L_2 = \prod_{t=2}^{T} \prod_{i=1}^{3} \prod_{j=1}^{3} q_{ij}^{Y_j^i(z_{t-1}, z_t, d)} P_j^{Y_j(z_{t-1}, z_t, d)}$$

Note that L_2 is a conditional likelihood given the selection in period one.

The likelihood function in the general case thus has the form

$$L = L_1 L_2 ,$$

which implies (11) of the text.

Consider next the parameters d_i defined on page 10 ($\lambda>0$). The definition of d_i, $i = 1, 2, 3, 4$, yields

$$\text{(A.1)} \qquad \begin{cases} d_2 = (3d_1 + d_4)/4, \\ d_3 = (d_1 + d_4)/4, \\ \quad d_1 < d_4, \end{cases}$$

and

$$\begin{cases} a_2 = 3(d_1 + d_4)^2/4 - 3(d_1 - d_4)^2/16, \\ a_1 = -3(d_1 + d_4)/2 . \end{cases}$$

This means that the parameters d_1 and d_4 determine the regimes as well as the coefficients a_1 and a_2. Moreover, if an observation z_t lies within (d_1, d_2) then the lower branch is selected whereas the middle branch is selected if $z_t \in (d_2, d_3)$, etc..

Let $\{d_{1m}\}$ and $\{d_{4n}\}$ be two finite sequences. For all m,n such that $d_{1m} < d_{4n}$ define a partition, P_{mn}, of this sample $\{z_t\}$ by

$$P_{mn} = \left\{(\min_t z_t, d_{1m}),\ (d_{1m}, d_{2mn}),\ (d_{2mn}, d_{3mn}),\ (d_{3mn}, d_{4n}),\ (d_{4n}, \max_t z_t)\right\}$$

where d_{2mn} and d_{3mn} are given by (A.1) with d_1 and d_4 replaced by d_{1n} and d_{4m}. Since the sample size is T the total number of partitions of the sample is T(T-1)/2. For a given partition a_1 and a_2 are given and the remaining parameters are to be determined by maximizing (12) and (13). Let the likelihood function that corresponds to these estimates conditional on P_{mn} be denoted by L_{mn}. The ML estimates are those which maximize L_{mn} with respect to (m, n).

Note that the standard errors reported above are only computed for the case with a unique equilibrium. To obtain the corresponding estimates in the case with multiple equilibria is possible through bootstrap simulation procedures, but we leave this for future work.