Tobin's q and Investment in a Model with Multiple Steady States

Mika Kato*Willi Semmler†and Marvin Ofori‡

May 2005

Abstract

This paper studies a simple dynamic investment decision problem of a firm where adjustment costs have capital size effects. This type of setting possibly results in multiple steady states thresholds and a discontinuous policy function. We study the global dynamic properties of the model by employing the Hamilton-Jacobi-Bellman method and dynamic programming that help us in the numerical detection of multiple equilibria and thresholds. We also explore the model's implications concerning the effects of aggregate demand, interest rates and tax rates. Finally, an empirical study on the firm size distribution is provided using U.S. firm size data. We utilize two different approaches, Kernel density estimation and Markov chain transition matrix to study an ergodic distribution. Our results suggest twin-peak distribution of firm size in the long run which empirically supports the theoretical conjecture of the existence of multiple steady states.

JEL Classification Numbers: E22
Key Words: Adjustment Costs, Multiple Steady States, Global Dynamics, Discontinuous Policy Function

*Department of Economics, Howard University, ASB-B-319, Washington, DC 20059. tel: (+1)(202)806-6774. fax: (+1)(202)806-5262. e-mail: mkato@Howard.edu.

†Center for Empirical Macroeconomics, Bielefeld and New School University, New York, 65 Fifth Ave, New York, NY 10003. tel (+1)(212)229-5783. fax:(+1)(212)229-5724. e-mail: semmlerw@newschool.edu.

‡University of Bielefeld, 33502 Bielefeld, Germany, e-mail:mofori@gmx.de

1 Introduction

In the theory of investment a firm maximizes the discounted value of its future net cash flows. What the firm wants to know is the optimal investment schedule over time. Generally speaking, by solving a firm's optimization problem, we want to obtain a policy function which tells us the optimal investment policy corresponding to each level of capital stock. This policy function has been perceived to be continuous. We, however, demonstrate the possibility of a discontnuous policy function in this paper. The key factor is to introduce adjustment costs with capital size effects, whereby the growth rate of capital is the source of adjustment costs. Interestingly, a recent work by Hartl, Kort, Feichtinger and Wirl (2000) provides a very simple investment model of this type. The purpose of our study is to solve a firm's dynamic investment decision problem where multiple steady states and a discontinuous investment strategy may arise. We study global dynamic properties of the model and pursue a numerical detection of the threshold.

It should be noted that economists have long been interested in studying the implication of adjustment costs on the dynamics of investment. An earlier version of such an investment study was presented by Jorgenson (1963). In his study, the optimal path of the capital stock for an exogenously given output is derived. The optimal path of investment rate is determined only when assuming distributed lag function. Those necessary ad hoc assumptions were recognized by Lucas (1967), Gould (1968), Uzawa (1968, 1969), and Treaday (1969). Their solution was to introduce the adjustment costs.

Generally speaking there are two types of adjustment costs: with and without capital size effects. Lucas (1967), Gould (1968) and Treadway (1969) developed the models with adjustment costs without size effects, whereas the models with adjustment costs with size effects have been studied by Uzawa (1968, 1969) and Hayashi (1982). Their works were stimulated by the earlier work on the adjustment costs by Eisner and Stroz (1963). Our model focuses on the role of adjustment costs with capital size effects. As it turns out, the capital size effects are central to the generation of multiple steady states and the associated itneresting dynamic properties. In addition to history dependence, both continuous and discontinuous policy functions can arise. Recently Hartl et al. (2000), using a simple framework with adjustment costs with size effects, discussed the case of multiple steady states. Our model in this paper is crucially based on their model but focuses more on the study of global dynamics using the HJB method, stresses the analysis of comparative dynamics and the policy implications, and explores the empirical support of the model's prediction on the firm size distribution.

The most important economic implication of the multiple steady states is that it has the property of history dependence. The history dependence arises if solution paths converge toward distinct attractors depending on the initial conditions. The existence of two or more stable steady states implies the existence of thresholds. At a threshold, the firm will be indifferent between an investment strategy converging toward one or the other stable steady states.

Hartl et al (2000) solve their model by using Pontriyagin's maximum principle and study the local properties of each steady state. We rather focus on comparing the values of the candidate paths and derive the global solution. We use the Hamiltonian-Jacobi-Bellman (HJB) method to derive a global value function and pave the way to detect thresholds numerically.

The remainder of the paper is organized as follows. Section 2 gives a brief review of the literature on adjustment costs in the investment theory. Section 3 presents a traditional model with adjustment costs with no capital size effects where we will obtain a unique and a continuous policy function. In other words, the optimal investment rate is continuous in Tobin's q. In Section 3, we present the model with capital size effects. Using the HJB method, the global value function and the threshold will be derived. The policy function can be both continuous and discontinuous. Numerical simulations are attached at the end of the section 3. Section 4 explores the model's implications concerning the effects of aggregate output (booms and recessions), interest rates and tax rates. Section 5 and 6 provide a empirical study on firm size distribution using the U.S. firm size data. We utilize two different approaches, Kernel density estimation and Markov chain transition matrix to obtain an ergodic distribution. Our results suggest twin-peak distribution of firm size in the long run which can be viewed as empirical supports of our theoretical model.

2 The Benchmark Model without Size Effects

We first present a traditional model without size effects. The model used here has been developed by Abel (1982), Hayashi (1982), and Summers (1981) and is known as the q theory model of investment. Generally speaking, this type of model has a unique equilibrium and a continuous policy function.

The model is represented in the equations (1)-(3):

$$\max_I \int_0^\infty e^{-rt}\{R(K) - I - A(I)\}dt \tag{1}$$

$$\text{s.t. } \dot{K} = I - \delta K;\ K_0 = \text{given} \tag{2}$$

where

$$R(0) = 0,\ R''(K) < 0,\ A(0) = 0,\ A'(I) > 0 \text{ for } I > 0,\ A'(I) < 0 \text{ for } I < 0,\ A''(I) > 0 \quad (3)$$

$R(K)$ is a representative firm's revenue function and $K(t)$ is a representative firm's capital stock. $I(t)$ is the firm's investment. We also assume the purchase price of a unit of investment goods is 1 and thus the cost of purchase investment goods is I. $A(I)$ represents simply adjustment costs which depends only on the firm's investment I. The depreciation rate of capital stock is δ .

For more specific results, we use the following specific functions which satisfy (3):

$$R(K) = aK - bK^2 \quad (4)$$

$$A(I) = cI^2 \quad (5)$$

We employ the Hamilton-Jacobi-Bellman (HJB) method to study the analytical solution of this problem. The HJB equation for the present model has the form:

$$rV(K) = \max_I [R(K) - I - A(I) + V'(K)(I - \delta K)] \quad (6)$$

Solving $\frac{\partial}{\partial I}[\cdot] = 0$ gives the first order condition:

$$-1 - A'(I) + V'(K) = 0. \quad (7)$$

Also, notice that $V'(K)$ is equivalent to Tobin's q. Therefore, we have the following rule for optimal investment:

$$I(t) = A'^{-1}[q-1]; \begin{cases} I > 0 & \text{for } q(t) > 1 \\ I = 0 & \text{for } q(t) = 1 \\ I < 0 & \text{for } q(t) < 1 \end{cases} \quad (8)$$

We can interpret q as the market value of a unit of capital. Since we assume that the purchase price of a unit of capital is 1, the economic interpretation of (8) is that a firm invests if the market value of capital exceeds the purchase price of capital and disinvest if the market value of capital is less than the purchase price of capital.

With specific functions, the optimal investment policy reads

$$I = \frac{1}{2c}(V'(K) - 1). \tag{9}$$

If e is a steady state level of capital stock then from (2),

$$I - \delta e = 0. \tag{10}$$

and, at the steady state, we obtain

$$rV(e) = R(e) - \delta e - A(\delta e) \tag{11}$$

$$V'(e) = \frac{R'(e)}{r + \delta}. \tag{12}$$

Thus, from the first order condition (7)and (2), the steady state e satisfies

$$-1 - A'(\delta e) + \frac{R'(e)}{r + \delta} = 0. \tag{13}$$

Using specific functions, (13) becomes

$$-1 - 2c\delta e + \frac{a - 2be}{r + \delta} = 0. \tag{14}$$

Therefore, we have a unique steady state:

$$e = \frac{a - (r + \delta)}{2c\delta(r + \delta) + 2b} \tag{15}$$

To establish a policy function, we construct a phase diagram in $\{I, K\}$ space. There exists a unique global stable manifold associated with a unique steady state. The global stable manifold in $\{I, K\}$ space is equivalent to the policy function and thus the policy function is continuous as Figure 1 shows.

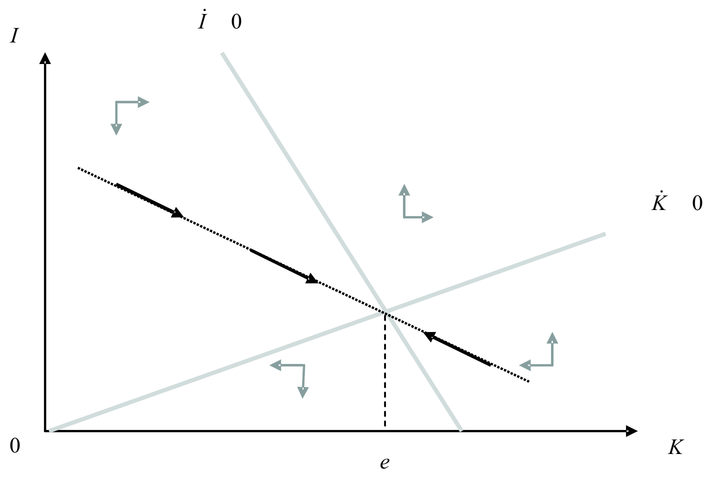

Figure 1: Continuous policy function

3 The Model with Adjustment Cost and Size Effects

We here follow the model developed by Hartl, et al (2000). The model is a simple dynamic investment model with relative adjustment costs with capital size effects. The simplest example is an adjustment cost function of investment to capital ratio as in Hartl et al. (2000). Multiple steady states and a discontinuous policy function may arise depending on parameters. We employ the Hamilton-Jacobi-Bellman method again to solve the problem with the aim to obtain a global value function. We want to detect the threshold numerically when we construct a policy function. The policy function can be both continuous and discontinuous depending on parameters.

Consider a firm acting to maximize the present value of the sum of future net cash flows.

$$\max_I \int_0^\infty e^{-rt}\left[R(K) - I - A\left(\frac{I}{K}\right)\right]dt \tag{16}$$

$$\text{s.t.} \quad \dot{K} = I - \delta K, \; K_0 = \text{given} \tag{17}$$

where $R(K)$ is a revenue function, K is a capital stock, I is investment, $A\left(\frac{I}{K}\right)$ is adjustment costs with size effects, i.e. the size of capital stock, K, affects adjustment costs, δ is a depreciation rate, and r is a discount rate. Here, by replacing the expression $\frac{I}{K}$ by u, our new control variable is u:

$$\max_u \int_0^\infty e^{-rt}[R(K) - uI(t) - A(u)]dt \tag{18}$$

$$\text{s.t.} \quad \dot{K}(t) = u(t)K(t) - \delta K(t), \quad K_0 = \text{ given.} \tag{19}$$

We assume that $R(K)$ and $A(u)$ are continuously twice-differentiable, and

$$R(0) = 0,\ R'(0) > 0,\ R''(K) < 0 \text{ for all} K \tag{20}$$

$$A(0) = 0,\ A'(0) = 0,\ A'(u) > 0 \text{ for } u > 0, \tag{21}$$
$$A'(u) < 0 \text{ for } u < 0,\ A''(u) > 0 \text{ for all } u.$$

We also assume that

$$R'(0) > r + \delta \tag{22}$$

so that the trivial solution $u = 0$ for all t is excluded.

For more specific results, let us assume that revenue and adjustment cost functions are quadratic:

$$R(K) = aK - bK^2 \tag{23}$$

$$A(u) = cu^2 \tag{24}$$

where $a, b, c > 0$.

The HJB equation for the present model has the form

$$rV(K) = \max_u [R(K) - uK - A(u) + V'(K)(uK - \delta K)]. \tag{25}$$

Step 1: Compute the steady states for the stationary HJB-equation.

Solving $\frac{d}{du}[\cdot] = 0$ gives the first order condition:

$$-K - A'(u) + V'(K)K = 0. \tag{26}$$

Again, it's important to notice that $V'(K)$ is equivalent to Tobin's q. The following rule of optimal investment is derived:

$$u(t) = A'^{-1}[(q-1)K]; \begin{cases} u > 0 & \text{for } q(t) > 1 \\ u = 0 & \text{for } q(t) = 1 \\ u < 0 & \text{for } q(t) < 1 \end{cases} \tag{27}$$

Comparing the investment rule under the relative adjustment cost with size effects (27) with (8), the rule under the adjustment cost with no size effects, the only difference is the capital stock K appears in the function u in a multiplicative way in (27). This implies that the firm with larger capital stock has higher incentive to invest and an increasing return to scale exists locally. This can be understood as the source of multiple steady states.

With specific functions (23) and (24), (27) can read

$$u = \frac{1}{2c}(V'(K) - 1)K. \tag{28}$$

If e is a steady state then from (19),

$$ue - \delta e = 0. \tag{29}$$

Since for any positive steady state $e > 0$, $u = \delta$,

$$rV(e) = R(e) - \delta e - A(\delta). \tag{30}$$

$$V'(e) = \frac{1}{r}[R'(e) - \delta]. \tag{31}$$

Thus, the first order condition (26) at the positive steady state $e > 0$ becomes

$$-e - A'(\delta) + \frac{1}{r}[R'(e) - \delta]e = 0. \tag{32}$$

From the specific functions,

$$R'(K) = a - 2bK \tag{33}$$

$$A'(u) = 2cu. \tag{34}$$

therefore, we obtain positive steady states from the condition (32):

$$-e - 2c\delta + \frac{1}{r}[a - 2be - \delta]e = 0. \tag{35}$$

Thus,

$$\begin{cases} u = 0 \text{ for } e = 0 \\ u = \delta \text{ for } e > 0 \end{cases} \tag{36}$$

and we have three steady states:

$$e = \begin{cases} 0 \\ \frac{a-r-\delta \pm \sqrt{(a-r-\delta)^2 - 16bc\delta r}}{4b}. \end{cases} \tag{37}$$

Note that two positive steady states exist if $a - r - \delta > 4\sqrt{bc\delta r}$.

Step 2: Solve the dynamic HJB-equation starting from the equilibrium candidates.

From the optimal investment rule (27), the satisfactory HJB-equation is

$$\begin{aligned} rV(K) \;&=\; R(K) - A^{'-1}((V'(K) - 1)K)K - \\ &\quad A(A^{'-1}((V'(K) - 1)K)) + V'(K)(A^{'-1}((V'(K) - 1)K)K - \delta K) \end{aligned} \tag{38}$$

For specific results, we substitute the optimal investment rule (28) and the satisfactory HJB-equation will be

$$V'(K)^2 - \frac{2K + 4\delta c}{K}V'(K) - \frac{4crV(K)}{K^2} + \frac{4ac}{K} - 4bc + 1 = 0. \tag{39}$$

Then we obtain an ordinary differential equation in V:

$$V'(K) = \frac{K + 2\delta c}{K} - \sqrt{\frac{(K + 2\delta c)^2}{K^2} + \frac{4crV(K)}{K^2} - \frac{4ac}{K} + 4bc - 1} \text{ for } K \geq e \tag{40}$$

$$V'(K) = \frac{K + 2\delta c}{K} + \sqrt{\frac{(K + 2\delta c)^2}{K^2} + \frac{4crV(K)}{K^2} - \frac{4ac}{K} + 4bc - 1} \text{ for } K < e \tag{41}$$

with

$$V(e) = \frac{1}{r}[ae - be^2 - \delta e - c\delta^2] \text{ for each } e > 0 \tag{42}$$

as initial conditions.[1]

Step 3: Solve the global value function.

We can compute the global value function for the original problem by

$$V(K) = \max V_i. \tag{43}$$

The local value functions V_i are generally computed numerically. The global value function is shown in Figure 2 corresponding to Figure 3, the phase diagram in $\{u, K\}$ space. This example shows the case when the middle unstable steady state has a focus property. The point where the local value function changes is called threshold or Skiba point. The phase diagram in $\{u, K\}$ space can also be expressed in $\{I, K\}$.

Using some numerical examples presented by Hartl et al. (2000), we carry out some simulations.

Case 1: (continuous policy function) the Skiba point coincides with the middle equilibrium candidate e_2.

Example:

$r = 0.3, \delta = 0.1, b = 0.6, c = 0.3, a = 0.74, K_2 = 0.07047, K_3 = 0.2129$

Case 2: (discontinuous policy function) the Skiba point does not coincide with the middle equilibrium candidate e_2.

Example:

$r = 0.2, \delta = 1.2, b = 0.6, c = 0.3, a = 1, K_2 = 0.017679, K_3 = 0.5665$, threshold$= 0.01$.

Once we have multiple steady states, there generally exist multiple local stable manifolds satisfying the system and the transversality condition. Our model has three steady states and there are two local stable manifolds corresponding to two stable steady states. Moreover, depending on the local dynamic property of the middle unstable steady state, these two stable manifolds overlap about the middle unstable steady state as Figure 3 shows. Our strategy to compare the values of those candidate paths is to use the global value function (43). Using the global value function shown in Figure 2, we can uniquely choose the optimal global stable manifold which generates

[1] As we see in (36) and (39), $u = 0$ for $e = 0$ and $u = \delta$ for $e > 0$. Thus, the value of the value function at each steady state is expressed as
$V(e) = \frac{1}{r}[ae - be^2 - \delta e - c\delta^2]$ for $e > 0$ and $V(0) = 0$ for $e = 0$.

the maximum value among the multiple candidate stable manifolds. The global stable manifold in $\{u, K\}$ space (or in $\{I, K\}$ space) is equivalent to the policy function and thus the policy function can be discontinuous as in Figure 3. Discontinuity will be easily obtained depending on parameters and we cannot rule out this case a priori. Note that the discontinuity occurs at threshold(s) where the local value function switches. For our parameter set above the threshold is found[2] to be at 0.01, thus to the left of the middle equilibrium candidate, K_2.

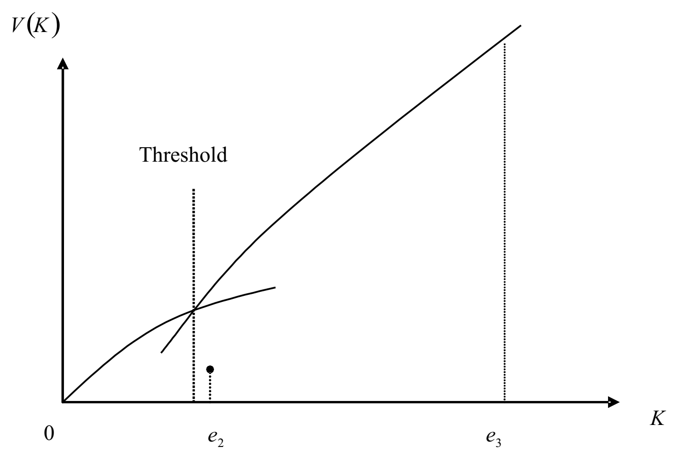

Figure 2: Global value function

[2]That threshold for the above parameter set is computed in Grüne and Semmler (2004).

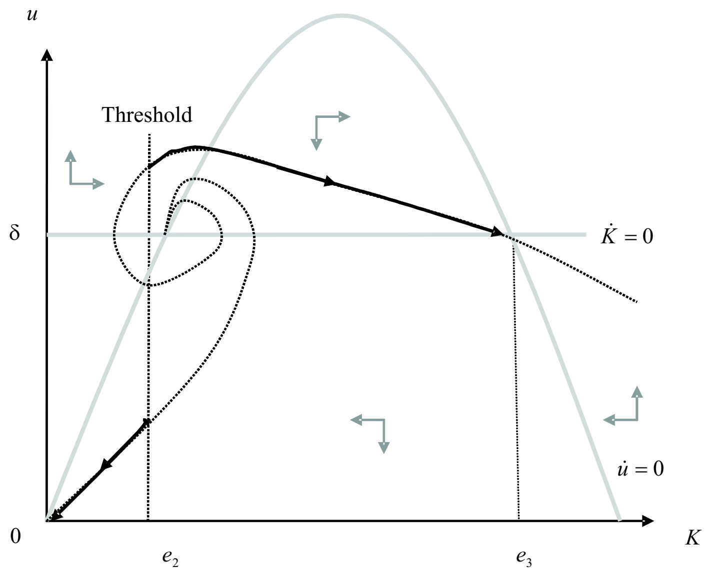

Figure 3: Discontinuous Policy Function in $\{u, K\}$ Space

4 The Effects of Demand, Interest Rate, and Tax Rate

In this section, we analyze the effects of exogenous changes in demand, interest rate, and taxes on the steady states and the equilibrium path. The exogenous increase in demand raises output of the firm and thus the revenue will increase for the same amount of capital, entailing an upward shift of the revenue function $R(K)$. This corresponds to an increase in a in our specific revenue function.

From (35), we have[3]

$$\frac{\partial e_2}{\partial a} = -\frac{e_2}{(a-r-\delta)-4be_2} < 0 \tag{44}$$

$$\frac{\partial e_3}{\partial a} = -\frac{e_3}{(a-r-\delta)-4be_3} > 0. \tag{45}$$

Output increase will make the middle steady state shifting down and the upper steady state shifting up.

[3]From (37), $e_2 = \frac{y-\sqrt{x}}{4b}$ and $e_3 = \frac{y+\sqrt{x}}{4b}$ where $y \equiv a-r-\delta > 0$ and $x \equiv (a-r-\delta)^2 - 16bc\delta r > 0$. Therefore, $e_2 < \frac{y}{4b}$ and $e_3 > \frac{y}{4b}$ hold. This implies the signs of (44) and (45).

In the same way, the effects of an increase in interest rate can be analyzed as

$$\frac{\partial e_2}{\partial r} = -\frac{e_2 + 2c\delta}{(a - r - \delta) - 4be_2} > 0 \tag{46}$$

$$\frac{\partial e_3}{\partial r} = -\frac{e_3 + 2c\delta}{(a - r - \delta) - 4be_3} < 0 \tag{47}$$

An increase in interest rate will shift the middle steady state up and the upper steady state down.

Concerning the effects of a tax rate, we introduce an investment tax credit for simplicity which is a direct rebate of χ percent of the price of capital. This changes the firm's optimal investment rule (27) as follows:

$$u(t) = A^{'-1}[(q + \chi - 1)]; \begin{cases} u > 0 & \text{for } q(t) + \chi > 1 \\ u = 0 & \text{for } q(t) + \chi = 1 \\ u < 0 & \text{for } q(t) + \chi < 1 \end{cases} \tag{48}$$

and three steady states become:

$$\tilde{e} = \begin{cases} 0 \\ \frac{a - r - \delta + \chi r \pm \sqrt{(a - r - \delta + \chi r)^2 - 16bc\delta r}}{4b} \end{cases} \tag{49}$$

Therefore, the effects of an increase in the rate of tax credit on positive steady states are:

$$\frac{\partial \tilde{e}_2}{\partial \chi} = -\frac{r\tilde{e}_2}{(a - r - \delta + \chi r) - 4b\tilde{e}_2} < 0. \tag{50}$$

$$\frac{\partial \tilde{e}_3}{\partial \chi} = -\frac{r\tilde{e}_3}{(a - r - \delta + \chi r) - 4b\tilde{e}_3} > 0. \tag{51}$$

An increase in the tax credit pushes the middle steady state down and the upper steady state up and thus, the domain of attraction associated with the upper steady state is enlarged.

Now our interest is how those changes in exogenous variables affect the firm's investment policy and its equilibrium path. It is convenient to construct the phase diagram for this purpose. We so far have relied on the method of HJB where the discussion can be boiled down to one dimensional state space. Here, we employ the Maximum Principle to see explicitly the equilibrium path in the control and state space, that is the equilibrium relationship between investment decision and the level of capital stock.

We follow the model (18)-(22) with the specific functions (23) and (24). The current value Hamiltonian[4] is

$$\begin{aligned} H &= R(K) - uK - A(u) + q\dot{K} \\ &= aK - bK^2 - uK - cu^2 + q(uK - \delta K) \end{aligned} \tag{52}$$

The first-order necessary conditions are

$$H_u = -K - 2cu + qK = 0 \tag{53}$$

$$\dot{q} = (r + \delta)q - (a - 2bK) + (1 - q)u \tag{54}$$

$$\dot{K} = (u - \delta)K. \tag{55}$$

Therefore, the system in u, K space is summarized as

$$\dot{u} = ru + \frac{1}{2c}(r + \delta - a)K + \frac{b}{c}K^2 \tag{56}$$

$$\dot{K} = (u - \delta)K. \tag{57}$$

Solving $\dot{u} = \dot{K} = 0$ gives the steady states which are same with (37).

[4]Introducing the tax credit χ will modify the current value Hamiltonian as: $H = aK - bK^2 - uK - cu^2 + (q + \chi)(uK - \delta K)$. The first order conditions are:

$$\begin{aligned} H_u &= -K - 2cu + (q + \chi)K = 0, \\ \dot{q} &= (r + \delta)q - (a - 2bK + \chi\delta) + (1 - q - \chi)u, \text{and} \\ \dot{K} &= (u - \delta)K. \end{aligned}$$

The system is rewritten as

$$\dot{u} = ru + \frac{1}{2c}(r + \delta - a - \chi r)K + \frac{b}{c}K^2 \text{ and}$$
$$\dot{K} = (u - \delta)K.$$

Solving $\dot{u} = \dot{K} = 0$ gives the steady states (49).

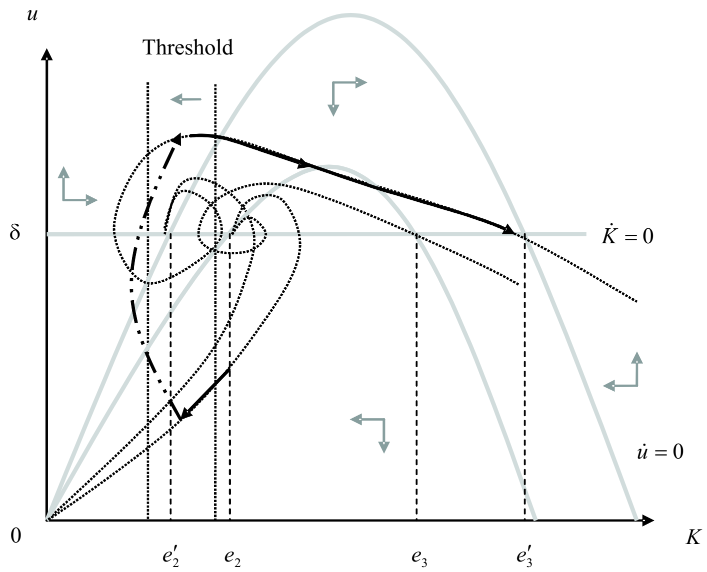

Figure 4: Comparative dynamic results

The phase diagram can be constructed by using the information from (52). We know from (44)-(47) and (50)-(51) that an increase in demand, a decrease in the interest rate, and an increase of tax credit shifts up the $\dot{u} = 0$ curve. The effects are shown in Figure 4. After the events occur, $\dot{u} = 0$ moves up and $\dot{K} = 0$ stays there. This changes the steady state values of the middle and upper steady states and, moreover, the threshold level. The threshold moves leftwards and therefore the domain of attraction changes, i.e. the domain of upper steady steady state's attraction enlarges. One of the interesting phenomena due to the system with multiple steady states is that such an exogenous change may bring us from one to the other trajectory, each of which is associated with a different steady state. For example, as Figure 4 shows, even though the firm's capital stock is on the shrinking trajectory at the beginning, the firm may switch to the growing trajectory after any of the suggested events, an increase in demand, a decrease in the interest rate or an increase of tax credit happens because of the leftward movement of the threshold. This is likely to happen when such an event occurs before the capital stock shrinks below the new threshold. As far as the market takes a turn for the better (an increase in demand) or the government takes a quick action to recover from a recession (through a decrease in the interest rate or an increase of tax credit) before things really get worse (before the firm's capital stock falls below the new threshold), there will be a chance for the firms to come back to the growing trend. It is also reasonable to think that firms will ride on the upswing of the market in a boom time, raising their investment. On the other hand, a recession may urge firms to make a change of direction, tightening their investment and their capital stock may even shrink. When the firm's investment is shrinking, not only the scale of policy intervention but also a quick reaction by the government is desired. When the government action is too late, a large policy intervention is required to alter the situation, since the government has to bring the threshold to the very low level. A quick reaction may be desired. The quicker the policy reaction is, the less costly and smaller intervention is required to save the situation.

5 Kernel Density and Boot Strap Test

Next we want to present empirical evidence that may confirm the long run twin-peak distribution of firm size which is implied by the above studied dynamic investment model. To address this question we concentrate on firm size distribution in the US-manufacturing industry.

The data of the following empirical study are taken from the *pstar* dataset used in Hall and Hall (1993). The data set contains 23 variables which quantify certain characteristics such as investment, stock price or assets' value of US-firms in the manufacturing industry for the time period 1960 to 1991. We use the variable *netcap* which is defined as book value of assets, adjusted for the effects of inflation to represent capital stock. In the following net capital which is normalized by the average net capital of all firms will be used as a measure of size:

$$k(i)_t = \frac{\text{Net Capital of Firm i in year t}}{\text{Average Net Capital in year t}} \tag{58}$$

Thus $k(i)_t = 1$ indicates that firm i has a net capital that equals the average net capital and $k(i)_t = 0.5$ means that firm i has half of the average net capital.

The pstar data set gives us a set of observed data points or capital stocks respectively which can be interpreted as a sample of an unknown probability density function for several years. To analyze certain characteristics of this density, such as the number of modes (which are defined as local maximums in the density), one has to determine the unknown density. If for example the density function has changed from being unimodal to a bimodal one it can be regarded as a hint for the conclusion of the above stated theoretical ideas.

Accordingly the aim in this section is to estimate the unknown density function from the observed data especially at the end of the sample period to see if it is bimodal. To address this aim a nonparametric approach is used. Whereas the parametric approach makes strong assumptions about the distribution of the data (e.g. normal distribution) the nonparametric approach is characterized by less rigid assumptions on the distribution of the observed data.[5] The only necessary assumption which will be made is that a probability density function exist, thus 'letting the data itself determine the density function'. In literature there are various kinds of nonparametric density estimators like kernels, splines, orthogonal series or histograms.[6] We will concentrate on one of the most common, namely kernel density estimation which recently has become a standard method in explorative data-analysis.[7]

The idea of kernel density estimators is based on Rosenblatt (1957) and Parzen (1962) and the simplest form is defined as

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} W(\frac{x - X_i}{h}) \tag{59}$$

where $x_i, (i = 1, \ldots, n)$ are the observations of the data set, n is the *window width* or equivalently denoted *smoothing parameter* or *bandwidth* and $W = W(u)$ represents the kernel. Let $f(\hat{x, h})$ be a kernel density estimate which uses data x and h as window width, then Figure 6 displays the number of modes in the Gaussian kernel density estimation in 1990 as a (step stair) function of the window width h, whereas the critical window widths are the points of discontinuity.

[5]See Silverman (1986), p. 1.

[6]For an detailed overview see for example Bean and Tsokos (1980) and Tapia and Thompson (1978).

[7]See Turlach (1993), p.1.

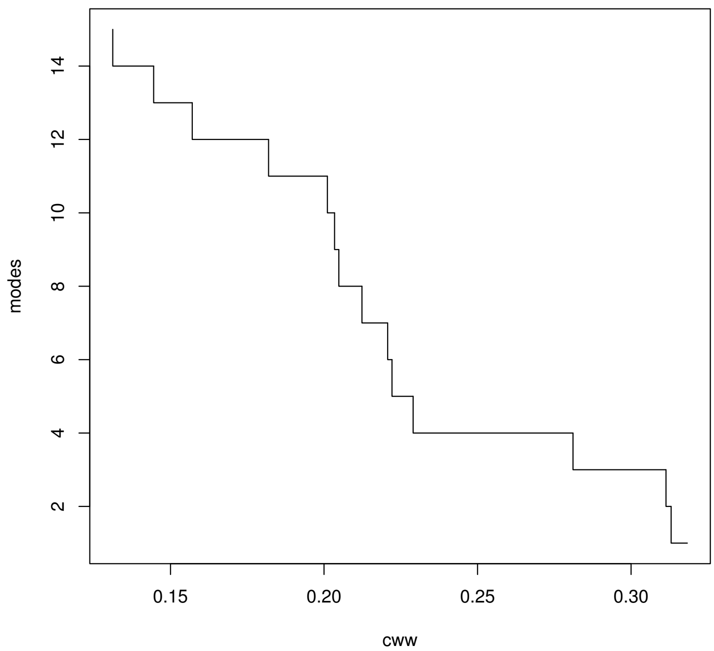

Figure 5: Number of Modes in the Gaussian Kernel Estimation in 1990

From Figure 5 it is obvious that inferences which can be drawn from the kernel estimator strongly depend on the choice of the window width h. Unfortunately yet there is no generally accepted approach to determine the "correct" window width. Recent approaches range from subjective choice to rather automatic ones, which try to define optimal criteria.[8]

Terrell (1990) assumes that the most part of density estimations are based on subjective choice, thus the researcher chooses the window width that best fits his aims. On this basis a lot of effort was undertaken to identify appropriate, data-driven window width selection methods.[9] The most popular choice for the window width is the so called *optimal bandwidth* introduced by Silverman (1986) and is defined as

$$h = 1.06\sigma n^{-1/5} \tag{60}$$

[8]The reader who is interested in a brief summary of these methods is referred to Turlach (1993).

[9]See for example Jones (1991), Silverman (1986) and many others.

The use of Silverman's optimal bandwidth for a Kernel estimate for the year 1990 results in an optimal bandwidth h=0.3183 which implies a kernel density estimation with one mode therefore indicating unimodality. This is shown in figure 6.

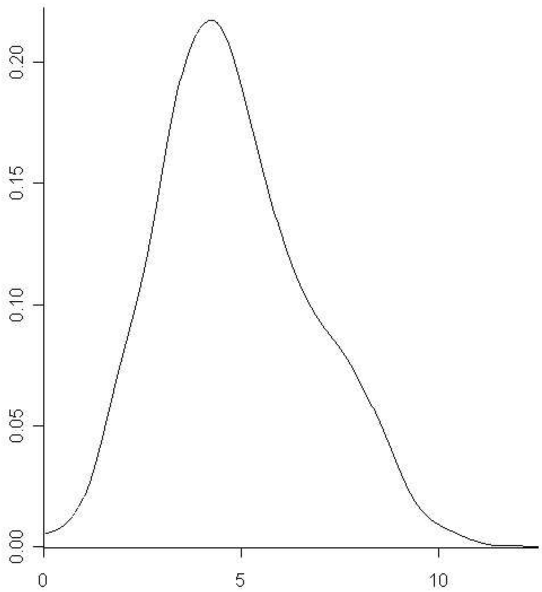

Figure 6: Kernel Density Estimate in 1990

Another way for addressing the problem of the number of modes is to perform hypothesis testing and therefore it is appropriate to construct a bootstrap test for multimodality which tests the hypothesis that a Gaussian density estimation has one, two or in general m modes. The approach of a bootstrap multimodality test is used because for example normal tests or permutation tests of multimodality are not efficient in this case. However, a Bootstrap test is an effective way to test for multimodality.

In the context of this work suppose that the data on firms' capital stocks for every year is an i.i.d. sample from an unknown distribution F with continuous density $p(x)$. To obtain certain properties of the density like the number of modes one can apply the bootstrap multimodality test which requires drawing bootstraps[10] from the empirical distribution F_n.[11]

Accordingly, the sample of firms' capital stocks can be used to represent the empirical distribution. The problem, considered in this section is testing the null hypothesis H_0 that the density has one single mode versus the alternative hypothesis H_1 that it has two or more modes.

$$H_0: \ f(x) \text{ has 1 mode} \quad \text{versus} \quad H_1: \ f(x) \text{ has more than 1 mode.}$$

To carry out this test it is necessary to specify a test statistic.
A reasonable choice is the critical window width $\hat{h}_1$ respectively the smallest window width needed to obtain a kernel density estimation with 1 mode. A large value of $\hat{h}_1$ implies that a lot of smoothing has to be done to obtain a kernel estimator with 1 mode and thus is evidence against H_0.

Next, one has to define the estimated distribution under the null hypothesis, respectively an estimated null distribution for the test of H_0. A reasonable choice may be $\hat{f}_0(x; \hat{h}_1)$ which intuitively is the density estimate that uses the smallest amount of smoothing among all estimates with one mode. The problem of using $\hat{f}_0(.; \hat{h}_1)$ is that it artificially increases the variance of the bootstrap sample relative to the variance of the actual data set. To avoid this problem $\hat{f}_0(.; \hat{h}_1)$ will be adjusted or rescaled respectively to have the same variance as the bootstrap sample variance and will be denoted smoothed density estimate $\hat{g}_0(.; \hat{h}_1)$.

Finally, it remains to assess the significance level of the observed value of $\hat{h}_1$. The bootstrap multimodality hypothesis test is based on the so called *achieved significance level* ASL_{boot}:

$$ASL_{boot} = P_{\hat{g}_0(.;\hat{h}_1)}\{\hat{h}_1^* > \hat{h}_1\}, \tag{61}$$

where $\hat{h}_1$ is fixed at its observed value from the data set and $\hat{h}_1^*$ is the critical window width producing one mode of the bootstrap sample x_i^*, $(i = 1, \ldots, n)$.

[10] Formally a bootstrap sample is constructed by randomly sampling n times with replacement from a data sample. The data sample or the distribution from which the bootstraps are drawn is denoted *empirical distribution* F_n and the replacement of F by F_n is called the "plug-in"-principle.

[11] It can be shown that F_n is a nonparametric maximum likelihood estimator of F and therefore it is justified to estimate the true unknown distribution by the empirical distribution if no other information on F is available (for instance such that F belongs to a specific parametric class).

To approximate the ASL_{boot} it is necessary to draw the bootstrap from the smoothed density estimate $\hat{g}_0(.;\hat{h}_1)$. The smoothed bootstrap data $x_i^*,\ (i = 1, 2, \ldots, n)$ will be obtained by drawing with replacement a sample $y_i^*,\ (i = 1, 2, \ldots, n)$ from the actual data set and then set

$$x_i^* = \overline{y}^* + (1 + h_1^2/\hat{\sigma^2})^{-1/2}(y_i^* - \overline{y}^* + h_1\epsilon_i),\ (i = 1, 2, \ldots, n), \qquad (62)$$

where $\overline{y}^*$ denotes the mean of $y_i^*,\ (i = 1, 2, \ldots, n)$, $\hat{\sigma^2}$ denotes the sample variance of $x_i^*,\ (i = 1, 2, \ldots, n)$ and ϵ_i are i.i.d. standard normal random variables. Now the ASL_{boot} is approximated by

$$A\hat{S}L_{boot} = \frac{\#\{\hat{h}_1^*(b) \geq \hat{h}_1\}}{B}, \qquad (63)$$

where B is the number of bootstrap replicates, # is the mathematical symbol for number and $(b = 1, 2, \ldots, B)$.

If for instance a significance level of 10% will be used, then H_0 will be rejected if $A\hat{S}L_{boot} < 10\%$.

After describing the basic ideas of the bootstrap multimodality test in the following it will be applied. Applying the bootstrap multimodality test

$$H_0:\ f(x) \text{ has 1 mode} \quad \text{vs.} \quad H_1:\ f(x) \text{ has two or more modes}$$

$$A\hat{S}L_{boot} = \tfrac{\#\{\hat{h}_1^*(b) \geq \hat{h}_1\}}{B},$$

with 500 smoothed bootstraps of the original data on logarithm of *netcap* in 1990, a significance level of 10% and $\hat{h}_1 = 0.35032$ results in 435 smoothed bootstrap samples that have $\hat{h}_1^*(b) \geq \hat{h}_1$ and therefore

$$A\hat{S}L_{boot} = \#\{\hat{h}_1^*(b) \geq \hat{h}_1\}/B = 87\%.$$

Analogous to the latter result the test indicates that the null hypothesis can not be rejected in favor of the hypothesis of a bimodal density distribution. The $A\hat{S}L_{boot}$ is about 87% which supports the hypothesis of just one single peak in the true density.

The main problem, however, using a (nonparametric) density estimator concerning long-run dynamics of firms is that it does not consider potential dynamics of firm distributions over time. In other words, although the true long-run distribution may indeed is bimodal, a density estimate in the year 1990 or even 2003 has not necessarily to be bimodal because it may take some long time for firms to group around two stable steady states. Therefore a density estimate that is unimodal for a certain time period does not contradict the hypothesis of a future (long-run) bimodal firm size distribution. To model the dynamics of firm size distribution and to determine long-run distributions that may be stationary (with two steady states) are studied in the next section. The Markov chain approach will be used to remedy the latter disadvantage.

6 Markov Chain Approach

Assume that the population of US-manufacturing firms is classified into several discrete size classes (states) according to some criteria that represent capital stock. In this work the value of a firm's book assets (*netcap*) in relation to the average value of all firms will be used, whereas other variables like investment, etc. could also be used, because in general it will be expected that they are strongly correlated. The *relative netcap* now is defined as

$$k_i(n) = \frac{K_i(n)}{\sum_i K_i(n)}, \tag{64}$$

where $k_i(n)$ denotes the relative netcap in year n, $K_i(n)$ is the absolute value of netcap in year n and the state space is considered as following:

$$0 \leq k \leq \tfrac{1}{4}, \qquad \tfrac{1}{4} < k \leq \tfrac{1}{2}, \qquad \tfrac{1}{2} < k \leq 1, \qquad 1 < k \leq 2, \qquad k > 2$$

The main problem of applying the Markov chain approach is the determination of the corresponding transition probabilities. It is obvious that the transition probabilities and the probability distribution respectively are not known and accordingly they have to be estimated from the observed data. To address this issue, one considers the variable "relative netcap" and interprets it as given realizations of a Markov chain. With the help of these realizations it is possible to determine the maximum likelihood estimator of the true transition probabilities. The procedure of obtaining the maximum likelihood estimates of the true transition probabilities begins with the determination of so called transition numbers. Transition numbers indicate the number of firms' transitions from state i to state j during two successive time points. The realizations of transition numbers can be summarized in form of a matrix, called fluctuation matrix $\mathbf{F}(n-1,n) = [f_{ij}(n-1,n)]_{i,j\in S,n\in\mathbb{N}}$, .

For purpose of illustration the fluctuation matrices for the time period 1973-74 is presented in table 1:

	$k \leq \frac{1}{4}$	$\frac{1}{4} < k \leq \frac{1}{2}$	$\frac{1}{2} < k \leq 1$	$1 < k \leq 2$	$k > 2$
$k \leq \frac{1}{4}$	837	12	0	0	0
$\frac{1}{4} < k \leq \frac{1}{2}$	12	177	12	0	0
$\frac{1}{2} < k \leq 1$	1	8	99	11	0
$1 < k \leq 2$	0	0	5	78	7
$k > 2$	0	0	0	2	129

Table 1: Example of a fluctuation matrix

To interpret these fluctuation matrices one might look, for example, at the first row of fluctuation matrix $\mathbf{F}(1973, 1974)$.

During the time period 1973 to 1974 837 firms that had a relative *netcap* corresponding to interval $k \leq \frac{1}{4}$ in 1973 remained there after one year. Analogous 12 firms that had a relative *netcap* corresponding to interval $k \leq \frac{1}{4}$ in 1973 transit to the interval $\frac{1}{4} < k \leq \frac{1}{2}$, and so on.

As in Onatski (2003) we used a sample period of 22 years (1973 - 1990) to allow the sample to be a balanced panel. Now it is possible to assess the maximum likelihood estimator. The maximum likelihood estimator for the true transition probabilities on basis of s realizations of a stationary Markov chain is given by:

$$\hat{p}_{ij} = \frac{\sum_{n=1}^{N} f_{ij}(n-1,n)}{\sum_{n=1}^{N} f_{i}(n-1,n)} = \frac{f_{ij}}{f_{i}\cdot} \tag{65}$$

It is worth noting that $\hat{p}_{ij}$ is the empirically derived relative frequency and therefore is only an approximation of the true transition probabilities. However, Anderson and Goodman (1957) have shown that this approximation is in deed a maximum likelihood estimate for a first-order Markov chain.[12]

Applying the maximum likelihood estimator results in the following transition matrix:

[12] See Anderson and Goodman (1957), pp. 90.

$$\mathbf{P} = \begin{pmatrix} 0.951 & 0.023 & 0.008 & 0.005 & 0.013 \\ 0.228 & 0.692 & 0.067 & 0.005 & 0.007 \\ 0.162 & 0.069 & 0.700 & 0.062 & 0.010 \\ 0.155 & 0.008 & 0.085 & 0.674 & 0.078 \\ 0.158 & 0.006 & 0.006 & 0.032 & 0.799 \end{pmatrix}$$

The estimated transition matrix **P** makes it possible to determine the equilibrium distribution or limiting distribution respectively of firms in the US-manufacturing industry. To guaranty that an equilibrium exist the transition matrix has to satisfy some properties which are presented in the following lemma:

Lemma:
An irreducible and aperiodic Markov chain with finite state space always has a limiting distribution which is an unique stationary distribution and furthermore is independent of the initial distribution of the Markov chain.

On the basis of above Lemma, first, it has to be proved that the transition matrix is irreducible and aperiodic:

1. A Markov chain is called *irreducible* if all states of the chain communicate. Thus, it must be possible to get from every state to every other in a finite number of steps. Respectively it must be satisfied that for each state pair $(i, j) \in S \times S$ exists an integer $n = n(i, j) \in \mathbb{N}$ with p_{ij}^{n}. Since every probability in the transition matrix **P** is strictly positive ($p_{ij}^{1} > 0$) it is obvious that the chain is irreducible.

2. Let a set U_i contain transition numbers indicating the number of steps, a state i returns to it with a positive probability. Then, if there is a state i such that $p_{ii}^{n} > 0$ for some $n > 0$ which states that there is a positive probability that the state i may be returned and furthermore the expression

$$d(i) = \begin{cases} \infty & \text{if } U_i = \{0\} \\ gcd\ of\ the\ set\ U_i & \text{if } U_i \neq 0 \end{cases}$$

equals 1 or ∞ then the state is denoted *aperiodic*. It is apparent that in this case the Markov chain is aperiodic.

After showing that the transition matrix, respectively the Markov chain, is irreducible and aperiodic there have to be a limiting distribution and a stationary distribution that is independent of an initial distribution. To determine the limiting distribution one has to solve the following equations:

$$\mathbf{p} = \mathbf{p}\mathbf{P} \tag{66}$$

$$\mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 + \mathbf{p}_4 + \mathbf{p}_5 = 1 \tag{67}$$

The solution of the equations results in a unique stationary distribution given by

$$\begin{aligned} \mathbf{p}_1 &= 0.787 \\ \mathbf{p}_2 &= 0.071 \\ \mathbf{p}_3 &= 0.046 \\ \mathbf{p}_4 &= 0.030 \\ \mathbf{p}_5 &= 0.066 \end{aligned}$$

It is obvious that this equilibrium distribution is bimodal or in the words of Quah[13] 'twin-peaked', since the probabilities of being in the extreme size groups in the equilibrium state distribution are larger then being in the middle state (which sometimes is referred to as thinning in the middle).[14]

In contrast, the results of the kernel density estimate indicated an unimodal distribution of US-manufacturing firms. However, the application of the Markov chain approach shows that in the long run a multiple steady state with a large group at the low tail and a smaller group at the high tail exists. Thus, the empirical study exhibits that in the long run the unimodal distribution ceteris paribus tends to change to a bimodal one. This implies an increase in the degree of concentration because there is a relatively small group (state 5: 7%) of firms having a capital stock that is at least twice as big as the average capital stock in the US-manufacturing industry, and a relatively large group of firms (state 1: 79%) that only have a capital stock corresponding to less then a quarter of the average capital stock in the industry.

[13] See e.g. Quah (1997), p. 28.

[14] See Quah (1993), p. 430.

With reference to the discrete Markov chain approach researcher sometimes argue that the choice of the class-size intervals has an important influence on the corresponding ergodic distribution. For instance, if one group is by accident divided into two pieces by using different definitions or numbers of class-size intervals, then the latter bimodal ergodic distribution may turn into a trimodal distribution, thus leading to a totally different interpretation of the equilibrium state.

A second problem associated with discrete Markov chains is that discretization can have great distortionary effects on dynamics if the underlying variables (which in our case is firm asset size) are continuous.[15]

However, just recently an approach called the stochastic kernel approach has been developed to overcome the shortcomings of discrete and arbitrarily defined class-size intervals. Informal, the idea of the stochastic kernel is to avoid the division into a discrete countable number of states and allowing the number of states tend to infinity and then to continuity. Thus, the resulting transition matrix becomes a matrix with a continuity of rows and columns or a stochastic kernel respectively. The dynamics of firms' asset size are assumed to be governed by this stochastic kernel similar to a first-order autoregression process.

To get an adequate understanding of the stochastic kernel consider the following: Let R denote the real-line and B the Borel σ-algebra on it. Furthermore let μ and ν elements that are probability measures on B. Then a *stochastic kernel* relating μ and ν is a mapping $M_{(\mu,\nu)} : (R, B) \to [0, 1]$ satisfying the three subsequent conditions:

1. $\forall x \in R$, the restriction $M_{(\mu,\nu)}(x, .)$ is a probability measure
2. For any $A \in B$, the restriction $M_{(\mu,\nu)}(., A)$ is B-measurable
3. For any $A \in B$, $\mu(A) = \int M_{(\mu,\nu)}(x, A)dv(x)$.

Though apparently the stochastic kernel is the uncountable generalization of a matrix or, in our case, of the transition matrix respectively. The stochastic kernel is completely sufficient to describe the transitions from state x to any other portion of the underlying state space.[16]

Next we will show of how the stochastic kernel can be used to model the dynamics of firms' asset sizes.

[15] See e.g. Chung (1974)
[16] See Quah (1997), p. 46.

By defining with λ_t the distribution of observations of firms' assets at time t, the stochastic kernel describes the law of motion as stated above through an operator M which maps the Cartesian product of the number of firms and the Borel set which can be measured on the space [0,1]. The simplest form of describing the law of motion is:[17]

$$\lambda_{t+1}(A) = \int M_t(x, A) d\lambda_t(x) \tag{68}$$

Accordingly, if one interprets $dv(x)$ as the density of λ and $d\mu(y)$ as the density of λ_{t+s}, then $M_{t+s}(x, dy)d\lambda_t(x)$ has to regarded as the conditional density of λ_{t+s} given λ_t. On the basis of this the stochastic kernel at time $t + s$ is equal to the transition matrix at $t + s$ with infinite uncountable numbers of rows and columns, thus each firms' asset size value represents its own class-size group. The problem of arbitrarily class-size intervals is solved because it is no longer necessary to construct artificial and subjective class-size intervals.

Again, using the balanced panel of the pstar data set for the period 1973-1990, the stochastic kernel is estimated. The stochastic kernel is estimated nonparametrically applying a Gaussian kernel and the optimal window width as suggested in Silverman (1986). To estimate the stochastic kernel, first the joint distribution of the logarithm of relative *netcap* is derived. Subsequently the implied marginal distribution in 1973 is determined by numerically integrating under this joint distribution. Finally, the stochastic kernel is obtained by dividing the joint distribution by the marginal distribution. The resulting stochastic kernel which relates the distribution of log firm assets in 1973 to the distribution of log firm assets in 1990 is shown in Figure 7.

[17]See Quah (1997), p. 47.

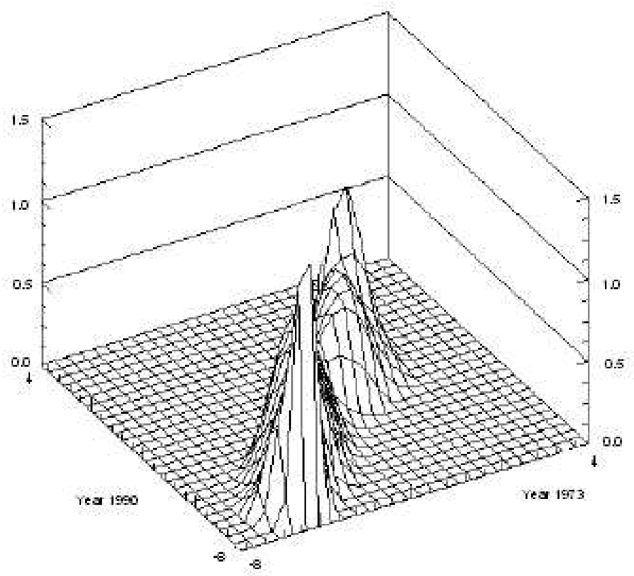

Figure 7: Stochastic Kernel of Log Firm Assets

The stochastic kernel estimate confirms the results of the transition matrix and ergodic distribution previously obtained. Figure 7 clearly indicates that the probability mass has two peaks at the two ends of the mass. As stated above this result gives empirical support of divergence and polarization with the tendency of the diverging states to cluster at high firm asset level, or at low firm asset level respectively.

7 Conclusion

In this paper we have studied a simple dynamic investment decision problem of a firm where adjustment costs have capital size effects. We have shown that with this type of models one can easily obtain in multiple steady states and thresholds as well as a discontinuous policy function, giving rise to a discontinuous behavior of investment. We study the global dynamic properties of such a model by employing the Hamilton-Jacobi-Bellman method and dynamic programming that help us in the numerical detection of multiple equilibria, the thresholds and the jump in investment. We also have explored the model's implications concerning the effects of aggregate demand, interest rates and tax rates. Finally, an empirical study on the firm size distribution is provided using U.S. firm size data. We utilize two different approaches, Kernel density estimation and Markov chain transition matrix to study an ergodic distribution. Our results suggest twin-peak distribution of firm size in the long run which can be viewed as empirically support of the existence of multiple steady states as predicted in the analytical part of the paper.

References

[1] Abel, A. B. 1982. "Dynamic Effects of Permanent and Temporary Tax Policies in a q Model of Investment," Journal of Monetary Economics 9: 353-373.

[2] Anderson, T.W. and L.A. Goodman. 1957."Statistical Inference about Markov Chains," Annals of Mathematical Statistics 28: 89-110.

[3] Bean, S.J. and C.P. Tsokos. 1980. " Developments in Nonparametric Density Estimation," International Statistical Review 48: 267-287.

[4] Boldrin, M., L. J. Christiano, and J. D. M. Fisher. 2001. "Habit Persistence, Asset Returns, and the Business Cycle," American Economic Review 91 (March): 149-166.

[5] Brock, W. A. and A. G. Malliaris. 1996. Differential Equations, Stability and Chaos in Dynamic economics. Amsterdam: North-Holland.

[6] Christiano, L., M. Eichenbaum, and C. Evans. 2001. "Nominal Rigidities and the Dynamic Effects of a Stock to Monetary policy," D'Autume, A. and P. Michel. 1985. " Future Investment Constraints Reduce Present Investment," Econometrica 53: 203-206.

[7] Chung, K.L. 1974. Elementary Probability Theory with Stochastic Processes. New York: Springer-Verlag.

[8] Deissenberg, C., G. Feichtinger, W. Semmler, and F. Wirl. 2004. "Multiple Equilibria, History Dependence, and Global Dynamics in Intertemporal Optimization Models" in: Economic Complexity: 91-122.

[9] Eisner, R and R. Stroz. 1963. Determinants of Business Investment, Impacts on Monetary Policy, Prentice Hall.

[10] Feichtinger, G., F. H. Hartl, P. Kort, and F. Wirl. 2000. "The Dynamics of a Simple Relative Adjustment-Cost Framework," German Economic Review 2(3): 255-268.

[11] Gould, J. P. 1968. "Adjustment Costs in the Theory of Investment of the Firm," Review of Economic Studies, 35:47-55.

[12] Grüne, L. and W. Semmler. 2004. "Using Dynamic Programming with Adaptive Grid Scheme for Optimal Control Problems in Economics", Journal of Economic Dynamics and Control, forthcoming.

[13] Hall, B.W. and R.E. Hall. 1993. "The Value and Performance of U.S. Corporations," Brookings Papers on Economic Activity 1: 1 - 49.

[14] Hartl, R. F., P. M. Kort, G. Feichtinger, and F. Wirl. 2000. "Multiple Equilibria and Thresholds due to Relative Investment Costs: Non-Concave - Concave, Focus - Node, Continuous - Discontinuous,"

[15] Hayashi, F. 1982. "Tobin's Marginal q and Average q: A Neoclassical Interpretation," Econometrica 50: 213-224.

[16] Jones, M.C. 1991. "The Roles of ISE and MISE in Density Estimation," Statistics & Probability Letters 12: 51-56.

[17] Jorgenson, S. and P. M. Kort. 1993. "Dynamic Investment Policy with Installation Experience Effects," Journal of Optimization Theory and Applications 77: 421-438.

[18] Kato M. and W. Semmler. 2001. "History Dependence and Hysteresis Effects in an Investment Model with Adjustment Costs,"

[19] Lucas, R. E. 1967. "Adjustment Costs and the Theory of Supply," Journal of Political Economy 75: 321-334.

[20] Mortenson, D. T. 1973. "Generalized Costs of Adjustment and Dynamic Factor Demand Theory," Econometrica 41: 657-667.

[21] Nickell, S. J. 1978. The Investment Decisions of Firms. Oxford: Cambridge University Press. Romer, D. 2001. Advanced Macroeconomis. McGraw-Hill Higher Education.

[22] Onatski, A. 2003. "Is the Firm Size Distribution Twin-Peaked?," Columbia University: Economics Department. Mimeo.

[23] Parzen, E. 1962. "On Estimation of a Probability Density and Mode," Annals of Mathematical Statistics 33: 1065- 1076.

[24] Quah, D.T. 1993. "Empirical Cross-section Dynamics in Economic Growth," European Economic Review 37: 426-434.

[25] Quah, D.T. 1997. " Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs," Journal of Economic Growth 2: 27-59.

[26] Rosenblatt, M. "Remarks on some Nonparametric Estimates of a Density Function," Annals of Mathematical Statistics 27: 823-837.

[27] Semmler, W. 1999. "Multiple Equilibria in Dynamic Economic Models,"

[28] Semmler, W. and M. Sieveking. 2000. "Global Dynamics in a Model with Multiple Equilibria,"

[29] Silverman, B.W. 1986. Density Estimation for Statistics and Data Analysis. London: Chapman and Hall.

[30] Skiba, A. K. 1978. "Optimal Growth with a Convex-Concave Production Function," Econometrica 46 (May): 527-539.

[31] Summers, L. H. 1981. "Taxation and Corporate Investment: A q-Theory Approach," Brookings Papers on Economic Activity no.1: 67-127.

[32] Takayama, A. 1985. Mathematical Economics. Cambridge: Cambridge University Press.

[33] Tapia, R.A. and J.R. Thompson. 1978. Nonparametric Density Estimation. Baltimore: John Hopkins University Press.

[34] Terrell, G.R. 1990. "The Maximal Smothing Principle in Density Estimation," Journal of the American Statistical Association 85: 470-477.

[35] Treadway, A. B. 1969. "On Rational Entrepreneurial Behavior and the Demand for Investment," Review of Economic Studies 36: 227-239.

[36] Tobin, J. 1969. "A General Equilibrium Approach to Monetary Theory," Journal of Money, Credit and Banking 1: 15-29.

[37] Turlach, B. A. 1993. "Bandwidth Selection in Kernel Density Estimation: a Review, " Working Papers in Statistic and Oekonometrie, Humboldt University Berlin

[38] Uzawa, H. 1968. "The Penrose Effect and Optimum Growth," Economic Studies Quarterly XIX: 1-14.

[39] .... 1969. "Time Preference and the Penrose Effect in a Two Class Model of Economic Growth," Journal of Political Economy 77: 628-652.

[40] Yoshikawa, H. 1980. "On the 'q' Theory of Investment," American Economic Review 70: 739-743.