Firm Dynamics with Customer Markets

Morten O. Ravn[*]
European University Institute and the CEPR

Stephanie Schmitt-Grohé[†]
Duke University, CEPR and the NBER

Martìn Uribe[‡]
Duke University and the NBER

May 21, 2007 -PRELIMINARY

Abstract

This paper introduces a customer markets model into a model of firm dynamics. We assume that household preferences are charaterized by goods-specific (deep) habits. This preference modeling implies that firms are faced with demand functions which relate their current demand to past sales. Firms are assumed to be monopolistically competitive and produce output using inputs of labor and are subject to idiosyncratic productivity shocks. There is free entry into the market but new entrants start their lifecycle with low levels of their customer stock. Incumbents leave the market when their continuation value are below their outside value. We solve for the stationary general equilibrium in which all markets clear and the expected value of entry is equal to zero. We show that this model gives rise to scale dependence in firm policies and in firm size dynamics. The key to the scale dependence is the induced dynamics in markups which are increasing in firm size and decreasing in productivity. The model is solved quantitatively. We find that customer markets can account for at least 1/3 of the growth difference between small and large firms.

Keywords: Customer markets, deep habits, firm dynamics, scale dependence, idiosyncratic productivity shocks

[*]Department of Economics, EUI, Villa San Paolo, via della Piazzola 43, FI-50133 Florence, Italy. Email: morten.ravn@eui.eu

[†]Department of Economics, Duke University. Email: grohe@duke.edu

[‡]Department of Economics, Duke University. Email: uribe@duke.edu

1 Introduction

This paper proposes customer markets as an important factor in accounting for firm dynamics. It is well-established that the firm size distribution is very unequal and approximately Pareto (Gibrat's law), and that small firms grow faster than big firms (scale dependence). A number of recent papers have suggested various mechanisms for accounting for these features such as industry specific human capital, selection, learning, or financial market imperfections.[1] We emphasize instead a demand side induced mechanism - a customer markets setting - which together with a selection process leads to substantial inequality in the firm size and to scale dependence in firm size dynamics. Thus, we highlight that imperfect competition structures together with customer markets effects might be important factors in accounting for firm dynamics. We also show that our theory has interesting and empirically relevant implications for a number of other cross-sectional features of firm variables.

We assume that firms are monopolistically competitive and are subject to idiosyncratic productivity shocks. Incumbent firms face fixed operational costs and entrants must pay a fixed entry cost to enter the market. The key aspect of our theory, however, is imbedded in the modeling of preferences which generate a customer markets model. We assume that households' preferences for individual goods varieties display habit persistence, the "deep habits" framework of Ravn, Schmitt-Grohe and Uribe (2006). According to this theory, households' demands for firms' products depend on the firms' past sales. This latter feature, which we refer to as a customer markets aspect, implies that firms face dynamic pricing problems. In particular, firms' need to take into account that their current prices affect the level of future demand for their products.[2]

The dynamic aspect of the demand facing firms has salient implications for the dynamics of mark-ups which have important repercussions on firm dynamics. We interpret the entry costs as investments in building up market research, building supplier and retailer networks, etc. which guarantees that the firm reaches a minimum level of demand. An entrant will therefore start its life with a minimum customer stock. Over time, the entrant can increase this customer stock, which leads to a higher level of demand, through increases in its past sales. Therefore, new entrants to the market will tend to set low levels of current mark-ups in order to build up their customer stock. This effect is stronger the more productive are the entrants

[1] Jovanovic (1982) assumes that firms learn about fixed costs so that unproductive slow growing firms gradually exit the market. Hopehayn (1992) and Ericson and Pakes (1995) generate selection effects through idiosyncratic productivity shocks. The importance of financial market frictions have been studied by Cooley and Quadrini (2001), Albuquerque and Hopenhayn (2004), and Clementi and Hopenhayn (2006).

[2] Other models of customer markets relationships include the switching cost theory, see Klemperer (1995), and informational theories, see Phelps and Winter (1970). We study the deep habits induced customer markets model because it allows for aggregation and is computationally much more convenient than alternative models.

since firms with low marginal costs can use the customer stock to insure themselves against the consequences of future adverse productivity shocks. This combined with selection implies that young firms will tend to have low markups and therefore grow fast. As firms mature, they increase their markups and this slows down their growth. This is the basic mechanisms that leads to scale dynamics in our theory.

The size dependence of the markup predicted by our theory is interesting. In particular, it is often claimed that large firms exploit market power to set higher markups and that this is harmful for welfare. While we do not explore the welfare effects of the presence of consumption externalities in our model, the size dependence of markups is not due to exploitation of market power. To be precise, the assumptions of constant price elasticity of aggregate demand and monopolistic competition implies that firms, as such, has no incentive to set higher prices as they grow bigger. Instead, the size dependence in markups is due to firms insuring themselves against future adverse productivity shocks when they are small. Thus, our theory offers an alternative view on the relationship between firm size and pricing behavior.

We show that the customer market mechanism is empirically relevant for firm dynamics and for firm size inequality. We do this by studying the properties of the stationary equilibrium of a calibrated version of the model. The stationary equilibrium displays firm turnover. Firms exit the market either because their initial productivity draw is low or because they are subject to a sequence of bad productivity shocks. Firms enter the market in the expectation of future profits. These mechanisms are common in models of firm dynamics, see e.g. Hopenhayn (1992). The customer markets feature of our model, however, propagates the idiosyncratic productivity shocks and introduces scale dependence. We show that the optimal policy functions are such that firms set markups that are increasing in their customer stock and decreasing in their level of productivity. Therefore, new entrants to the market that choose not to exit initially set low markups and typically have high levels of productivity. Therefore, these firms tend to grow fast. As the firms mature having built up their customer stock. A large customer stock allows mature firms to survive extended periods of low productivity. Therefore, larger (and older) firms will exit the market less frequently than small (and young) firms.

Quantitatively, we find that the growth rate of small firms is around 5 percent higher than the growth rate of large firms. This is a sizeable effect comparable to the quantitative role of other mechanisms of scale dependence studied in the literature. Moreover, the stationary distribution of employment across firms displays substantial amount of inequality of firm sizes. We show that the resulting firm size distribution approximates a Pareto distribution quite well. Therefore we claim that imperfectly competitive market structures combined with customer markets features such as our deep habits formulation might be important for understanding the dynamics of firms and the size distribution of firms. We also show that it is the propagation of selection through customer markets that is key for our results.

The remainder of the paper is structured as follows. The next section describes the economic environment and defines the stationary equilibrium Section 3 discusses the calibration and quantitative analysis of the model. Section 4 contains the main results. Section 5 discusses implications for asset pricing. Finally, we conclude in Section 6.

2 The Model

We analyze firm dynamics in a dynamic stochastic general equilibrium model of imperfect competition. There is a large number of monopolistically competitive firms that produce a differentiated good using inputs of labor. Labor is assumed to be perfectly mobile across firms. Firms' levels of productivity are subject to idiosyncratic productivity shocks that are assumed to be persistent. There is free entry into the industry and production is subject to fixed operational costs. New entrants to the industry pay a fixed entry costs and incumbents are free to leave the market whenever they should desire to do so.

A key aspect of our theory is the specification of the demand side of the economy. We introduce a customer markets aspect by assuming the existence of deep habits as in Ravn, Schmitt-Grohe and Uribe (2006). According to this theory, households' preferences for the differentiated goods produced by the firms are characterized by habit formation at the level of individual goods. We adopt an external habit formation model since this simplifies the analysis very considerably compared to an internal habit model.[3] This model of customer markets implies that firms are faced with demand functions in which the location of their current demand depend positively upon their past sales. Firms therefore face dynamic pricing problems and set markups that depend both upon their current (and future expected) productivity level and upon their past sales. This latter aspect is the central new element of our analysis and we will show that it has interesting and relevant implications for firm dynamics.

We first describe the household side of the economy and then turn to the analysis of firms. After that we define the equilibrium concept of the economy and derive expressions for key variables that we later characterize quantitatively. The quantitative evaluation of the theory is performed in the preceding section.

[3] Ravn, Schmitt-Grohe and Uribe (2006) show that firms' pricing policies are subject to time-inconsistency problems when deep habits are internal. Nakamura and Steinsson (2006) analyze a sustainable plans equilibrium for an economy with deep internal habits. These papers derive symmetric equilibria for economies with aggregate risk. We focus instead on idiosyncratic productivity risk in which case the model becomes very complicated when allowing for internal habits. Moreover, we believe that the external habits formulation has intuitive appeal in that the habits capture effects due to market penetration, marketing efforts, brand recognition, etc. that are external to individual consumers.

2.1 Households

Unlike many models of firm dynamics, the demand side is key for our analysis. We introduce a customer markets feature into households' demand for firms products. The supply side repercussion of this aspect of the theory is central for understanding why we can account for size dependency in the exit rates of firms. The demand side is therefore what distinguishes our theory from other models of size-dependent firm turnover.

There is a continuum of identical infinitely lived households of measure 1 indexed by $j \in [0;1]$. Each household supplies inelastically one unit of labor and labor can move freely across firms. We will abstract from population growth as it has no major implications for our theory. Households consume a basket of goods defined over a continuum of individual goods c_{it}^{j} for $i \in [0; I_t]$ where I_t denotes the measure of varieties available at date t.

Household j's preferences are described by:

$$U_t^j = E_t \sum_{s=t}^{\infty} \beta^{s-t} \log X_t^j \tag{1}$$

where β is the discount factor of the household and E_t denotes the mathematical expectations operator conditional on all information available at date t. X_t^j denotes a bundle of "habit adjusted" consumption goods defined by:

$$X_t^j = \left(\int_0^{I_t} \left(\frac{c_{it}^j}{s_{it}^{\theta}} \right)^{1-1/\eta} di \right)^{1/(1-1/\eta)} \quad , \ \theta\left(1-1/\eta\right) \geq 0 \tag{2}$$

where s_{it} is defined as:

$$s_{it} = \{ \begin{matrix} (1-\rho)\, s_{it-1} + \rho c_{it-1} \text{ if variety } i \text{ existed in period } t-1 \\ s_0 \text{ if variety } i \text{ did not exist in period } t-1 \end{matrix} \tag{3}$$

In this equation $c_{it} = \int_0^1 c_{it}^j dj$ denotes the average cross-sectional consumption of good variety i. Our specification of the intertemporal utility function implies that the intertemporal elasticity of habit adjusted consumption (but not of consumption itself) is equal to unity. This assumption has no consequence for our results.

The aggregator X_t^j is assumed to have constant elasticity of substitution over habit adjusted consumption levels of the goods, see equation (2). The elasticity of substitution is given by η and we will assume that $\eta > 1$. The CES structure is very convenient computationally and has been adopted in almost all previous contributions to the literature. Moreover, combined with our modeling of habits, the specification of X_t^j implies that the price elasticity of demand is constant, a property that is important for the dynamics of mark-ups.

According to equation (2), the marginal utility of household j's consumption of an *individual* goods variety i depends on the past cross-sectional average consumption of this variety. This formulation corresponds to the "relative deep habits" model of Ravn, Schmitt-Grohe and Uribe (2006).[4] The parameter θ measures the strength of this habit effect and $\rho \in (0,1)$ is the rate of depreciation of the habit stock. When $\rho = 1$, the habit stock is given by the past cross-sectional consumption of good variety i while values of $0 < \rho < 1$ implies that the habit stock depends on all past consumption levels of the good.

In the following we will assume that $\eta > 1$ and that $\theta \leq 0$. When $\theta = 0$, the model implies that current marginal utility of consumption is independent of past consumption. When $\theta < 0$, $\partial X_t^j / \partial c_{it}^j$ is increasing in s_{it} so that the current marginal utility of consumption of this variety of any given individual household is increasing in the past aggregate consumption of the good.

For any given level of X^j, household's demand for individual goods varieties must solve the cost minimization problem:

$$\min \int_0^{I_t} p_i c_i^j di \tag{4}$$

$$s.t. \quad \left(\int_0^{I_t} \left(\frac{c_i^j}{s_i^\theta} \right)^{1-1/\eta} di \right)^{1/(1-1/\eta)} \geq X^j \tag{5}$$

where p_{it} denotes the price of goods variety i in period t. Solving this problem gives the demand functions:

$$c_i^j = \left(\frac{p_i}{P} \right)^{-\eta} s_i^{\theta(1-\eta)} X^j \tag{6}$$

where $P = \left(\int_i s_i^{\theta(1-\eta)} p_i^{1-\eta} di \right)^{1/(1-\eta)}$ is an aggregate price index.

The interesting aspect of the demand function in equation (6) is that each household's demand for good variety i depends positively on the past cross-sectional average consumption of this good as summarized by s_i. In other words, a firm with higher past sales will, ceteris paribus, be faced with a higher level of current demand. This is a key implication akin to other customer markets models in which past sales affect firms' current the customer stocks, see e.g. Klemperer (1995) for a model with brand-switching costs or Phelps and Winter (1970) for a model with information based customer market effects. The advantage of our theory is that it leads to customer market effects while preserving tractability of the demand side of the economy. In particular, our theory allows for aggregation, an aspect that simplifies very considerable the analysis when there is idiosyncratic risk on the supply side. We will refer to s_{it} as the customer stock (or as the habit stock).

[4] Ravn, Schmitt-Grohe, and Uribe (2006) concentrate on the case with "difference" habits. This formulation is very convenient in models with aggregate shocks only as it gives rise to a strong tendency for counter-cyclical mark-ups in response to demand shocks. The reason is that demand functions under this specification display increasing price elasticity in response to increases in aggregate demand. The relative habits assumed in the current paper instead leads to constant price elasticity demand functions.

Households own the firms in the economy and we assume that households have access to a complete set of contingent claims markets. Each household chooses its demand for X_t^j and its portfolio of contingent claims by maximizing (1) subject to a sequence of budget constraints:

$$X_t^j + E_t r_{t,t+1} d_{t+1}^j = d_t^j + w_t + \Psi_t^j$$

where $E_t r_{t,t+1} d_{t+1}^j$ is the period$-t$ price of a random payment d_{t+1}^j in period $t+1$, w_t is the period t real wage defined as the nominal wage divided by the price index P_t and Ψ_t^j is the real dividend income on the household's holdings of equity. Real dividend income includes firms' transfers of profits to the household sector as well as transfers from the household sector to firms that may be making losses. We assume that households hold a diversified portfolio consisting of claims to the net current profits of the entire set of currently active firms in the economy.

We can then formulate the households' intertemporal maximization problem as:

$$\begin{aligned} V^h\left(d^j\right) &= \max\left[\ln X^j + \beta E V^h\left(d^{j\prime}\right)\right] \\ s.t. &: X^j + Er' d^{j\prime} = d^j + w + \Psi^j \\ d^{j\prime} &\geq d_{\min} \end{aligned}$$

where we let $d^{j\prime}$ denote next period's asset holdings of household j and $V^h\left(d^j\right)$ is the optimal value of the consumers' objective function given its asset holdings. $d_{\min}$ denotes a borrowing limit which we will assume is exogenous.

The model of firms that we describe below assumes that there are only idiosyncratic productivity shocks. Moreover, we will focus upon the properties of the stationary equilibrium of the economy. For these reasons, the main insights from the household sector of the economy are summarized by the demand functions in equation (6), the law of motion for the habit stocks given by equation (3), and by the households' exogenous supply of labor. Aggregating across households gives us that these quantities are given as:

$$c_i = \left(\frac{p_i}{P}\right)^{-\eta} s_i^{\theta(1-\eta)} X \tag{7}$$

$$s_i = \begin{cases} (1-\rho)\, s_{i,-1} + \rho c_{i,-1} \text{ if variety } i \text{ existed in period } -1 \\ s_0 \text{ if variety } i \text{ did not exist in period } -1 \end{cases} \tag{8}$$

$$n = \int_j n^j dj = 1 \tag{9}$$

$$X = \int_0^1 X^j dj \tag{10}$$

where n denotes aggregate labor supply, n^j is household j's labor supply, and X is a measure of aggregate goods demand. As we mentioned earlier, a convenient feature of our customer markets model is that it allows for aggregation while at the same time capturing the essential implications of customer markets models.

2.2 Firms

There is a continuum of incumbent firms of measure I and a continuum of potential entrants. Firms produce a differentiated good and act as monopolistically competitors in the output market while taken prices for given in the input markets. In order to produce, firms must pay a fixed operational cost. Entrants must pay a fixed entry cost prior to entering the market. The existence of fixed operational costs and fixed entry costs implies that there is entry and exit in the equilibrium. We assume that output is produced using a single variable input, labor.

We follow earlier contributions to the firm dynamics literature and assume that firms are subject to *idiosyncratic* productivity shocks, c.f. Hopenhayn (1992), Ericson and Pakes (1995), or Luttmer (2004). Aggregate productivity is assumed constant. Given the demand structure discussed above, firms are characterized by their current productivity and by their customer stock as indicated by their habit stock, s_i. This latter aspect adds a new dimension to firm dynamics which is at the centre of our analysis.

In order to produce, an incumbent firm must pay a fixed (nominal) operational cost, $\varphi > 0$. Having paid this cost, firms produce output using inputs of labor. Labor is assumed to be a homogenous good. The production function is specified as:

$$y_{it} = A_{it} n_{it}^{\alpha} \tag{11}$$

where A_{it} denotes firm i's level of productivity in period t. This variable is assumed to be stochastic. In particular, we assume that A_{it} follows an M-state discrete valued homogeneous Markov process, $A_{it} \in \mathcal{A} = [a_1, ..a_M]$, $a_1 > a_2 > .. > a_M$. The transition probability matrix is given by the the matrix Q with elements $q_{kl} = \Pr(A_{it+1} = a_k | A_{it} = a_l)$. We let Q_∞ with elements q_k^∞ denote the invariant distribution associated with Q.

The current period real profits, π_i, of firm i are given as:

$$\pi_i = \frac{p_i}{P} c_i - w \left(\frac{c_i}{A_i} \right)^{1/\alpha} - \varphi \tag{12}$$

Given the firm's choice of p_i and its productivity, its labor demand follows as:

$$n^d(s_i, A_i) = \left(\frac{c_i}{A_i} \right)^{1/\alpha} \tag{13}$$

where we suppress the dependence of labor demand on real wages and on the level of aggregate demand, X.

Incumbent firms can at any point in time decide to leave the market. As in Hopenhayn (1992), we assume that an incumbent firm that wishes to leave the market in period t must make its exit decision at the end of period $t-1$ prior to observing next period's productivity shock. Said differently, the fixed operational costs are paid (or has to be committed to) prior to observing the current period productivity. An incumbent firm's maximization problem can be formulated recursively as:

$$
\begin{aligned}
V\left(s_i, A_i = a_k\right) &= \max_{p_i, c_i, s_i'} \left[\pi_i + \beta \max\left(\sum_{h=1}^{M} q_{hk}(V\left(s_i', A_i' = a_h\right), 0\right)\right] \qquad (14) \\
s_i' &= \left(1-\rho\right) s_i + \rho c_i \\
\pi_i &= \frac{p_i}{P} c_i - w\left(\frac{c_i}{A_i}\right)^{1/\alpha} - \varphi \\
c_i &= \left(\frac{p_i}{P}\right)^{-\eta} s_i^{\theta(1-\eta)} X
\end{aligned}
$$

The formulation of the incumbents' problem assumes a fixed discount factor, β, which coincides with households subjective discount factor. This reflects two aspects of our analysis. First, we assume that firms are owned by the household sector. Secondly, we will focus upon the properties of stationary equilibria in which aggregate variables are constant.

An incumbent's problem is dynamic for two reasons. First, the existence of fixed operational costs means that the firm faces an exit decision which depends on its expectations regarding A_i' given A_i and on its future customer stock. When expected future productivity is low and it has a low customer stock, a firm might decide to exit the market because the expected present value of future revenues are too low relative to variable and fixed costs. This is reflected in the inner maximum operator in the value function in equation (14): whenever the expected value of remaining in the market is non-positive, the firm will exit and its value from then on is equal to the outside option which we normalize to 0.

Secondly, whenever $\theta < 0$, firms' current pricing choices affect their future profit opportunities. A firm with a larger customer stock will be better equipped to survive periods of high marginal costs since the level of demand facing the firm is increasing in s_i. Higher past sales, as reflected in s_i, effectively shifts the demand curve out. For a mature firm, this effect is not very important. For a young firm, however, current sales are likely to be large relative to the customer stock and the effect of current sales on future demand is therefore relatively large for such firms. A young (and growing) firm may therefore set low current prices of its good in order to insure itself against low future productivity. A combination of a low customer stock and low productivity, however, makes it likely that the firm chooses to exit the market. As the firm matures, the insurance value of its customer stock becomes less important.

When instead $\theta = 0$, incumbent firms will set prices as a constant mark-up, $\mu = \eta/\left(\eta - 1\right)$, over marginal costs and its exit decision is entirely dictated by its expectations regarding future productivity (relative to fixed operational costs).

Ravn, Schmitt-Grohe and Uribe (2006) examine the dynamics of mark-ups under a deep habits specification. That paper shows that goods-specific habits have a strong tendency for generating countercyclical mark-ups in response to demand shocks. There are several differences between the current economy and the one examined by Ravn, Schmitt-Grohe and Uribe (2006). First, we focus upon the effects of idiosyncratic productivity shocks while Ravn, Schmitt-Grohe and Uribe (2006) focus upon aggregate shocks. Secondly, Ravn, Schmitt-Grohe and Uribe (2006) assume homogeneous firms and derive the properties of symmetric equilibria. In the present economy, firm's choice of prices will be symmetric only to the extent that two firms are characterized by identical state variables, (s_i, A_i). Third, the current paper assumes relative habits while Ravn, Schmitt-Grohe and Uribe (2006) for much of their analysis focus upon difference habits.[5] This aspect is important. The relative habit formulation adopted in the present paper implies that dynamics in firms' mark-ups are due only to intertemporal effects. In the difference habit specification instead, mark-ups vary over time due to both the intertemporal effects and due to changes in price elasticities when current demand changes relative to habitual demand. While the difference habit specification is well-suited to environments with aggregate risk (and in which one can focus on symmetric equilibria), the relative habit formulation is better suited to individual firm pricing problems.[6] As we shall show later, the current model has interesting and plausible implications for the cross-section of mark-ups across firms. Finally, our focus are upon the dynamics follow the effects of productivity shocks rather than demand shocks.

At any point in time, there is a continuum of potential entrants to the market. At the beginning of the period, prior to observing their productivity, potential entrants decide whether or not to enter the market. Entry requires paying a fixed (nominal) entry cost $\Phi > 0$. If an entrant pays this fee, it receives an initial "market share" $s_{i0} > 0$ and draws an initial productivity shock from $\mathcal{A}$. It is assumed that the probability of drawing $A_{i0} = a_h$ is given by the ergodic probability of a_h, i.e. by q_h^{∞}. s_{i0} indicates the location of the demand facing a new firm that enters the market. We assume that $s_{i0} = s_0$ so that new firms are homogenous in terms of their initial customer stock. We make this assumption because we will think of s_0 as a minimum level of customers that is required for a firm to guarantee a positive amount of sales. In this setting Φ can thus be interpreted as the amount of market research, marketing activities, costs of building supplier and retailer networks such that the firm can sustain its minimum customer market stock.

[5] Ravn, Schmitt-Grohe and Uribe (2007) also adopt relative habits in an examination of cost pass-through.

[6] In the difference habit formulation, the demand facing individual firms contains a term that is completely price inelastic. Therefore, firms perceivably have an incentive to set infinitely high prices for one period only. Such pricing behavior can be ruled out in symmetric equilibria if there is only aggregate risk. In a heterogenous firm setting it is harder to rule out such behavior. This incentive does not arise in the relative habit formulation as demand tends to zero when prices go to infinity.

The problem of an entrant can be formulated recursively as:

$$\begin{aligned} V^{entry} &= \max_{p_i, s_i'} [\sum_{h=1}^{M} q_h^{\infty} \pi \left(A_i = a_h\right)_i^{entry} \\ &+\beta \max\left(\sum_{h=1}^{M} q_h^{\infty} \sum_{d=1}^{M} q_{dh} (V\left(s_i', A_i' = a_d\right), 0 \right)] \end{aligned} \tag{15}$$

where:

$$\begin{aligned} \pi\left(A_i\right)_i^{entry} &= \frac{p_i}{P} c_i^{entry} - w\left(\frac{c_i^{entry}}{A_i}\right)^{1/\alpha} - \varphi \\ c_i^{entry} &= \left(\frac{p_i}{P}\right)^{-\eta} s_0^{\theta(1-\eta)} X \\ s_i' &= (1-\rho)\, s_i + \rho c_i^{entry} \end{aligned}$$

V^{entry} in equation (15) denotes the value of entry gross of fixed entry costs. Firms will thus enter the market as long as $V^{entry} \geq \Phi$. Note that, since the entry decision is made prior to the entrant knowing its productivity level, the profits in the first period are stochastic with the probability of each state being given by q_h^{∞} (this is also reflected in the definition of the continuation value). An entrant that decides to remain in the market will face the dynamic problem of an incumbent from the next period onwards. Given a value of s_0, the entrant's exit decision is dictated by its realized productivity A_i. Therefore, entrants that draw sufficiently high levels of initial productivity will remain in the market while entrants with low productivity levels will immediately exit the market.

Our formulation assumes that all entrants to the market enter the market with identical customer stocks, s_0. It is straightforward to allow for heterogeneity in s_0 but we have not explored the consequences of this generalization.

Firms are owned by households and we assume that there are no financing constraints (apart from those that might be due to households' borrowing constraints). Thus, the owners of the firms, i.e. the households, will transfers funds to entrants in order to cover the fixed entry costs and to incumbent firms that are making current losses. The latter possibility can arise since the future expected value of firms might be positive even if current profits are negative. Indeed, as we will see later, small entrants with high productivity which have a high expected present value of future profits (and therefore a high expected value), may find it optimal to go through a period of substantial losses.

Alternatively, one might assume financing constraints which limit the losses that firms can make in any period. Such financing constraints will imply that firms might wish to retain earnings in order to insure themselves against future low levels of productivity and has consequences for the extent to which profit opportunities allow new firms to enter the market. Moreover, in the face of imperfect information about relevant firm characteristics, financial markets may have important consequences for the shape of firm dynamics, c.f. Cooley and Quadrini (2001), Albuquerque and Hopenhayn (2004) or Clementi and Hopenhayn (2006). We will neglect such features in the present analysis in order to focus entirely upon the effects of customer markets on firm dynamics.

Before continuing we find it instructive to gain some intuition for the dynamics of the model by inspecting the implications of deep habits for markup variations. We consider the problem of an incumbent and we simplify the setting by assuming that the firm never exits the market. In this case, the optimal markup must satisfy the following condition (see Appendix 1 for a derivation):

$$\omega_{it} = \left(\frac{\eta}{\eta-1}\right)\left(1+\theta\rho\frac{\eta-1}{\eta}\beta E_t\sum_{j=1}^{\infty}\beta^{j-1}\left(1-\rho\right)^{j-1}\omega_{i,t+j}\frac{mc_{i,t+j}}{mc_{i,t}}\frac{c_{i,t+j}}{s_{i,t+j}}\right) \tag{16}$$

where $\omega_{it} = p_{it}/\left(P_t mc_t\right)$ denotes the markup of firm i in period t, $mc_{i,t}$ denotes the marginal costs of firm i, and $E_t x_{t+j}$ denotes the expectation of x_{t+j} given all information available at date t.

Note that in the absence of habits, $\theta = 0$, the markup is constant and equal to $\left(\frac{\eta}{\eta-1}\right)$. When $\theta < 0$, the above condition implies that when current marginal costs are low relative to expected future marginal costs, the firm sets a low current markup. The reason for this is that a low current markup helps the firm build up its customer stock and the "price" of this is low when marginal costs are low. Moreover, a low customer stock implies a low markup. The latter insight is interesting for our purposes since "young" firms (firms that recently entered) are characterized by low customer stocks. Therefore, young firms have, ceteris paribus, a tendency for setting low markups thereby giving rise to size dependency in the firm dynamics.

2.3 Equilibrium

We study the properties of the stationary equilibrium of the model. Given that there are no aggregate shocks, the assumption of complete markets implies that aggregate quantities and prices will be constant in the stationary equilibrium. The stationary equilibrium, however, displays firm dynamics due to the presence of idiosyncratic risk. We will first define some objects of interest and we will then define the stationary equilibrium which properties we investigate in the preceding section.

We let Ω denote the set of (s, A) for which incumbent firms find it optimal to exit the market. Let $H \geq 0$ denote the measure of entrants and let I be the total measure of firms active in the market. Free entry implies that $V^{entry} \leq \Phi$, with strict equality if $H > 0$. Note from the incumbents' value function that the present discounted value of an incumbent's profits is increasing in A_i and in s_i. Fixed operational costs then imply that there exists a continuation rule such that the firm will exit the market the first time that $(s_i, A_i) \in \Omega$. This set of critical values of the firms' state variables is defined as:

$$\Omega = \inf \left\{ (s, A) \in \mathcal{S}x\mathcal{A} : \beta \sum_h q_{hk}(V\left(s', A' = a_h\right) | \left(s, A = a_k\right) = 0 \right\} \tag{17}$$

where $\mathcal{S} = [\underline{s}, \overline{s}]$ denotes the set of permissible values for s_i. Thus, the set Ω contains the combinations of current productivity and the customer markets indicator, s, such that a firm in the market decides to leave the market. Given weak conditions on the parameters of the model, this set will be non-empty and this guarantees that a strictly positive measure of firms leave the market.

Let z denote the probability distribution over s_i and A_i for entrant firms. Given our assumptions regarding the initial productivity draw and the initial level of s_{i0}, this probability distribution is given as:

$$z\left(s_i, A_j\right) = \begin{cases} q_j^{\infty} \text{ for } s_i = s_{i0} \\ 0 \text{ otherwise} \end{cases} \tag{18}$$

In words, all entrants start with the same level of $s_{i0} = s_0$, and the initial productivity level is a stochastic variable with a probability distribution that is drawn from the invariant distribution of Q.

We define $\mu\left(s, A\right)$ as the joint distribution of firms over s and A. In order to simplify notation, but also preempting our numerical analysis, we assume that s is discrete valued. More precisely, we assume that $s \in \mathcal{S} = [s_1, .., s_K]$ where $0 < s_1 < s_2 < .. < s_K$. Given the entry and exit decisions and incumbent firms' policy functions for s_i', which we denote by $S\left(s_i, A_i\right)$, the distribution of firms over $\left(s_i, A_i\right)$ evolves as:

$$\mu'\left(s', A'\right) = \sum_A \sum_{[s:s'=S(s,A)]} I\left(s, A\right) Q\left(A'|A\right) \mu\left(s, A\right) + H'z\left(S\left(A, s_0\right), A'\right) \tag{19}$$

where $I\left(s, A\right)$ is an indicator function that takes the values 0 if $\left(s, A\right) \in \Omega$ and 1 otherwise. The first term on the right hand side of (19) is the joint distribution of $\left(s', A'\right)$ for the incumbent firms that decide not to exit. $H'z\left(s', A'\right)$ is the corresponding measure for the entrants.

With these definitions at hand, we can now define the stationary equilibrium of the economy:

Definition 1 *A stationary monopolistic competition equilibrium with free entry is consists of policy functions $S\left(s, A\right)$, $y\left(s, A\right)$, $p\left(s, A\right)$, $n\left(s, A\right)$, a set Ω, a probability distribution $\mu\left(s, A\right)$, and positive real numbers $\left(P, w, X\right)$ such that:*

(i)The policy functions $S\left(s, A\right)$, $y\left(s, A\right), p\left(s, A\right)$, and $n\left(s, A\right)$ solve the firms' problems,
(ii) Firms exit when $\left(s_i, A_i\right) \in \Omega = \inf \left\{ (s, A) \in \mathcal{S}x\mathcal{A} : \beta \sum_{A'} Q\left(A'|A\right) \left(V\left(s', A'\right) | \left(s, A\right) = 0\right\}\right.$

(iii) The value of entry equals the cost of entry $V^{entry} = \Phi$ *if* $H > 0$, *otherwise* $V^{entry} < \Phi$
(iv) Households maximize their utility and observe their budget constraints
(v) All markets clear:

$$\begin{aligned} y(s,A) &= \left(\frac{p(s,A)}{P}\right)^{-\eta} s^{\theta(1-\eta)} X \\ X &= \left(\sum_A \sum_s \left(\frac{y(s,A)}{s^\theta}\right)^{1-1/\eta} \mu(s,A)\right)^{1/(1-1/\eta)} \\ 1 &= \sum_A \sum_s n(s,A)\,\mu(s,A) \\ P &= \left(\sum_A \sum_s s^{\theta(1-\eta)} p(s,A)^{1-\eta}\,\mu(s,A)\right)^{1/(1-\eta)} \end{aligned}$$

(vi) The distribution $\mu(s,A)$ *is a stationary distribution associated with* $S(s,A)$:

$$\mu(s',A') = \sum_A \sum_{[s:s'=S(s,A)]} I(s,A)\,Q(A'|A)\,\mu(s,A) + H'z(S(A,s_0),A')$$

In the stationary equilibrium aggregate variables are constant but idiosyncratic productivity shocks creates heterogeneity in individual firm dynamics. It is the features of the firm dynamics that we will now analyze. Given the complexity of the firms' problem, we will investigate the properties of the stationary equilibrium numerically.

3 Quantitative Analysis

In this section we wish to evaluate the extent to which our customer markets theory has important quantitative implications for firm dynamics in equilibrium. We initially focus upon the consequences for the size distribution for firms and for the size dependency of firm growth rates. However, since our modeling of customer markets, as we shall see, also has interesting implications for the dynamics of mark-ups, we will also investigate the implications for price dynamics.

Naturally, our analysis abstracts from many other features that most likely are important for the understand firm dynamics. Amongst the features that we exclude are capital accumulation, learning about fixed costs (Jovanovich, 1982), firm and industry specific human capital (Rossi-Hansberg and Wright, 2007), research and development in new product lines (Kortum and Klette, 2003), financial market frictions (Cooley and Quadrini, 2001, Albuquerque and Hopenhayn, 2004, Clementi and Hopenhayn, 2006), and heterogeneous managerial ability (Lucas, 1978, Garicano and

Rossi-Hansberg, 2004). Therefore, the aim of our analysis is mainly to analyze the extent to which the customer markets effects are quantitatively relevant rather than to see whether we can account fully for the firm size inequality and for the scale whole extent of scale dependency in firm size dynamics.

3.1 Calibration and Solution Procedure

We calibrate the economy so that each period relates to one year. We assume that $\beta = 0.96$ which corresponds to an annual real interest rate of 4 percent. This is a standard estimate in the literature.

Key parameters in our analysis pertain to the process of the idiosyncratic productivity process. Unfortunately, there is substantial uncertainty surrounding the calibration of this process.[7] Cooper, Haltiwanger and Willis (2004), estimate idiosyncratic shock processes based on quarterly employment and hours plant level data contained in the Longitudinal Research Database. They use a simulated method of moments estimator for a labor adjustment cost economy, and assume that the productivity shock process is given by first-order autoregressive process, $\ln A_{it} = \lambda \ln A_{it-1} + \varepsilon_{it}$ where $\varepsilon_{it} \sim nid\,(0, \sigma^2)$. Depending upon the specification of the labor adjustment costs, their estimates of λ vary from 0.180 to 0.984 and their estimates of σ are between 0.77 and 0.04.[8] Clearly, these estimates are enormously different. Moreover, the implied cross-sectional variance of total factor productivity are order of magnitude higher than the estimates of e.g. Comin and Philippon (2004). We follow Khan and Thomas (2007) and set $\lambda = 0.859$, but we assume a variance of the innovation that is somewhat higher than their estimate (and closer to Copper, Haltiwanger and Willis, 2004). We assume that $\sigma^2 = 0.04^2$.[9] This calibration implies that the ratio of the highest value of A to the lowest value of A is approximately 1.6 which implies a relative modest marginal cost difference. Given the uncertainty towards the calibration of these parameters, we also look at the sensitivity of the results to the persistence and variance of the technology shock process. We approximate the autoregressive processes with 9 state discrete valued first order Markov processes following the procedure in Tauchen (1986).

[7] Given the uncertainty towards the calibration of this process, Khan and Thomas (2006) simply assume that the idiosyncratic technology shock process follows the same process as standard estimates of the aggregate technology shock process.

[8] At the quarterly frequency, they find $\phi_q = 0.996$ and $\sigma_q = 0.01$ when assuming quadratic costs. When assuming fixed adjustment costs instead, they find that $\phi_q = 0.65$ and $\sigma_q = 0.33$. to translate these estimates into the annual frequency we assume that $\phi_a = \phi_q^4$ and that $\sigma_a = \left(1 + \phi_q + \phi_q^2 + \phi_q^3\right)^{1/2} \sigma_q$ (we use "q" to denote the quarterly frequency and "a" to denote the annual frequency). These approximations assume that productivity is end-of-period sampled and interprets this variable as a stock variable.

[9] Khan and Thomas (2007) assume that $\sigma^2 = 0.022^2$. With such a low value of the innovation variance, the difference between the lowest and the highest value of the idiosyncrtic productivity shock is so low that there is almost no entry and exit in equilibrium.

Next we need to calibrate the costs of entry and the fixed operation costs. According to the World Bank (2007), a narrow measure of resources spent on entry correspond to around 0.7 percent of GNI in the United States. It is hard to evaluate whether this value underestimates or overestimates true entry costs. While one might be skeptical about the World Bank estimate, we use it in the calibration due to the lack of better estimates. Thus, we calibrate Φ so that total entry costs in the economy correspond to 0.7 percent of aggregate value added. This implies a value of Φ that corresponds to approximately 15 percent of the value of the smallest entrant that chooses to remain in the market.

The fixed operational costs, ϕ, are calibrated such that the firm exit rate equals 13 percent per year. This estimate agrees with Davis, Haltiwanger and Jarmin (2006) and is also in line with the evidence in Rossi-Hansberg and Wright (2007). This gives us an estimate of ϕ that corresponds to approximately 1 percent of the nominal sales of the median firm in the stationary distribution.

The calibration of ϕ determines the fixed costs component of the firms total costs. The variable cost component is determined by α. Our theory makes no distinction between firms and plants and excludes other factor inputs than labor. Therefore, we will assume that $\alpha = 1$ which implies constant marginal costs.[10]

The last set of parameters that need to be calibrated relate to preferences and include η, the price elasticity of demand, and the habit parameters θ, the strength of habits, and ρ, the persistence of the habit stock. In the marketing literature, estimates of η are typically in the neighborhood of 2. Such a low estimate of the degree of substitutability between goods is a problem for standard theories of monopolistic competition because the implied markup is implausible high (100 percent). Estimates of markups instead imply much higher values of η. Morrison (1990), for example, estimate markups of around 1.2 which suggests that η is around 6. Our theory can potentially help reconciling these estimates since the markup depends not only on η but also on the habit parameters as well as on other parameters that affect the stationary distribution of firms.

We take the following approach. We assume that $\eta = 4$ which is in-between the values typically estimated in the marketing literature and the values consistent with the evidence on markups. We set $\rho = 0.25$ which is close to the annual value implied by the estimates of Ravn, Schmitt-Grohe and Uribe (2006).[11] We then calibrate θ so that the implied markup across firms is close to 10 percent. This is a low target but consistent with the values estimated by e.g. Morrison (1990). We set a relatively low target for the markup for one particular reason. As we will show in the next section, the markups are strongly scale dependent. In particular, larger firms typically set

[10] See Rossi-Hansberg and Wright (2007).

[11] Ravn, Schmitt-Grohe and Uribe (2006) estimate the persistence of habits to be around 0.86 in quarterly data. Their estimate, however, relates to a model with aggregate shocks and symmetric firms. The value of ρ that we assume implies a quarterly persistence of approximately 0.7 which is still relatively high.

higher markups. Selection bias is therefore likely to lead to an upward bias in the estimate of markups. The resulting estimate of θ is 0.4 which implies an average markup of 9.6 percent in the stationary equilibrium.[12]

We solve the model numerically using discrete state-space stochastic value function iterations. We discretize $\mathcal{S}$ (the space of the stock of habits) assuming that it consists of 300 equally spaced points between s_0 and $s_{\max}$. The lower limit of $\mathcal{S}$ is chosen to coincide with the initial customer stock of entrants. This assumption can be generalized but we find it natural given that s_0 has the interpretation of the investment that needs to be incurred in order to build up a minimum customer stock. $s_{\max}$ has to be chosen so that it is strictly inside the ergodic equilibrium distribution of the economy. Accordingly, we choose a value that is marginally higher than the customer stock associated with a firm which has an infinite sequence of productivity shocks equal to the highest value of A.

The value function iterations are straightforward to compute since the model is recursive. We use bi-section on the wage rate and on the number of firms in order to reassure that the labor market clears and that the value of entry equals the fixed entry cost. Walras' law then guarantees that the output market clears. This algorithm always converged in the experiments that we carried out, a feature that indicates the existence and uniqueness of the stationary equilibrium.

4 Results

In this section we examine the quantitative implications of our theory. Our main interest is in investigating the extent to which the customer markets model may help accounting for scale dependence in the firm dynamics. The model obviously abstracts from many other relevant features and therefore we do not expect it to account accurately for all the relevant features of the firm dynamics.

4.1 Firm Values and Policy Functions

Figure 1 illustrates the value of the firm for each level of productivity graphed against the level of the customer stock, s. The value of the firm is concave in s and increasing in the level of productivity. In this figure we also illustrate the locus of Ω, the set of (s, A) pairs for which an existing firm wishes to exit the market. The value of small and unproductive firms is non-positive and they leave the market. Large firms remain in the market even if they are hit by very bad productivity shocks since the demand for their products, thanks to their large customer stock, remains high. This dynamics is clearer when Figure 1 is combined with Figure 2 which shows the policy function for s_i'. Together with the policy functions we also plot a 45 degree

[12] This estimate is close to Martins et al's, 1996, estimates of markups in the US manufacturing sector.

line (along which $s'_i = s_i$). Firms with the highest level of productivity grow until they reach the rest point s_1^* (this is indicated with the full-drawn black line). This is the maximum customer stock size in the ergodic set of the stationary equilibrium. Firms with the next highest level of productivity grow until they reach the rest point s_2^* (also shown with a full drawn black line). Firms with low productivity instead decline fast and eventually leave the market (when the continuation value hits the zero bound). A firm that is born with the highest level of productivity and remains highly productive, reaches s_1^* in around 20 years. This event, however, occurs with a very low probability. Should such a firm suddenly find itself with the lowest level of productivity, it is able to survive for around 7 years before it eventually closes down. Thus, there is substantial persistence in the induced firm dynamics.

Figures 3 and 4 gives the details on the firms' pricing policies. In Figure 3 we show the (relative) price set by the firms and in Figure 4 we show the implied markup policies. The latter are computed as $p(s,A) / \left(\frac{w}{\alpha A} h(s,A)^{1-\alpha}\right)$ where w denotes the equilibrium wage and $h(s,A)$ is the employment policy function. The price policy functions have two important features. First, more productive firms set lower prices. This is natural since these firms produce with lower marginal costs. Such a feature would arise naturally in almost all models in which firms have some degree of monopoly power. In the customer market setting, this feature of the pricing policy implies that more productive firms will tend to increase their customer stock over time, a feature that is evident from the policy function for $s'(s,A)$. Figure 4 shows that the lower prices set by more productive firms is not only due to lower marginal costs but also due to lower markups. Effectively, very productive firms invest in market share in order to insure themselves against adverse changes in their marginal costs. This is the key feature of our customer market model.

Secondly, Figures 3 and 4 show that, conditionally on the level of productivity, larger firms set higher prices and higher markups. Recall that due to our assumption of constant marginal costs, size dependence in prices (for a given level of productivity) can be due only to size dependence in markups. Our model has exactly this feature. Holding constant the level of productivity, firms increase their markup as they mature. There are two reasons for this. First, as firms get bigger, their incentives to charge low markups due to their effect on the level of future demand disappear. Effectively, when the habit stock is large it becomes to expensive for the firm to invest in its customer stock. This can be seen from equation (3) - when s_i is large, the firm must sell a lot in order to maintain its customer stock and this comes at the cost of a low markup. Secondly, the insurance value of the customer stock becomes - in relatively terms - smaller when the firm gets very big. More precisely, if s_i is large, the firm can survive for an extended period should it be subject to an adverse productivity shock while a small customer stock implies that the firm goes out of business very fast should it experience very low productivity. Thus, the tendency to set low markups in order to increase the customer stock is mostly relevant for small firms.

One very interesting aspect of the implied markup dynamics is that small but productive firms set markups that are smaller than one. In other words, their revenues do not cover the variable costs. This behavior is due to the investment value of the customer stock being very high for these firms. Small but not very productive firms, instead, set high markups since it is too expensive for them to invest in the customer stock.

Figures 3 and 4, however, also show that there is no straightforward *unconditional* correlation between firm size and markups (and prices). Large and highly productive firms may or may not set higher prices than small and unproductive firms. In our calibration, the unconditional correlation between firm size and markups is actually negative since large firms are predominantly very productive while small firms are a mix of productive (and growing firms) that set low prices and unproductive firms that either have just entered the market or are about to exit.

Finally, in Figures 5 and 6 we show the resulting policy functions for $y(s, A)$ and $n(s, A)$. Both output and employment are strongly scale dependent and sensitive to the level of productivity. Firms with a higher level of the customer stock produce more output and employ more workers. Since these firms also are more productive, this implies that there is a small share of large productive firms with a large customer stock. In practise, given the process for the idiosyncratic productivity shock, there is only a small share of firms with the highest level of productivity but a substantial amount of firms with the second and third highest level of productivity. These firms have around 1500-5000 times as many employees as the smallest firms in the stationary equilibrium. Moreover, given that prices are decreasing in productivity, more productive firms employ more labor and produce disproportionately more output than less productive firms. Therefore, the distribution of employment and output is more unequal in our model than in models with competition or models with constant markups.

4.2 Firm Distribution and Growth Dynamics

We now examine the implications that our theory has for the size distribution of firms and for the dynamics of firm size. A large literature has documented that there is substantial inequality in the firm size distribution and that it follows, approximately, a Pareto distribution, a finding known as Gibrat's law (see Sutton, 1997, for a recent survey of theory and evidence on Gibrat's law). Our theory is not designed specifically to generate a Pareto distribution so we do not see this as a litmus test of its plausibility.

Scale dependence in firm dynamics has been documented in a number of papers and most recently by e.g. Rossi-Hansberg and Wright (2007). These authors study longitudinal data on establishment growth rates and on establishment entry and exit rates collected by the U.S. Census Bureau in the Business Information Tracking System (BITS).[13] They confirm earlier findings in the literature. In particular, smaller

[13] Unfortunately, this dataset is not publicly available.

firms grow faster than larger firms. According to the BITS data, firms with less than 10 employees grow with around 5-20 percent per year while firms with more than 10 employees have little or no employment growth on a year-to-year basis. Moreover, this growth differential persists over longer horizons. Exit rates are also strongly scale dependent with firms with less than 10 employees displaying exit rates around 16 percent per year while large firms having exit rates between 2 and 4 percent per annum. Thus, in summary, smaller firms (that survive) grow faster but there are also many small firms that exit the market.

The stationary distribution of s is illustrated in Figure 7. The limits of this distribution is the ergodic set. The lower of these limits is partially dictated by our assumptions regarding s_0 while the upper limit is determined endogenously. The distribution is very unequal.[14] There is a large number of small firms and a small, but significant amount of very large firms. Thus, in terms of the habit stock, our theory produces substantial amounts of inequality across firms.

Figure 8 illustrates the stationary distribution of employment. In order to compute this graph we defined a grid for employment and we then computed the density of firm employment values within the cells of this grid. The grid consists of 100 equally sized values distributed equally between the minimum and the maximum value of employment in the stationary distribution. In Figure 8A we graph the density of the employment distribution together with a Pareto distribution (with parameter 1) indicated by the dotted line in the diagram. The employment distribution produced by our model produces substantial inequality and the fit to the Pareto distribution is relatively good. The latter is somewhat surprising given that we made no special assumptions aimed at producing this result. It is clear that there are more small to medium sized firms in the stationary distribution than predicted by the Pareto distribution. The same, however, is true in US data (see Figure 1 in Rossi-Hansberg and Wright, 2007). We view the implications of our theory for firm size inequality as an indication that it is well-suited to account for this aspect of firm dynamics.

The relationship between firm size and firm growth rates is depicted in Figure 9. Consistently with the empirical evidence, our theory implies scale dependence in the sense that small firms grow faster than big firms. According to our calibration, the growth rate of small firms is approximately 5 percent while larger firms do not grow and the largest firms display negative growth rates. This estimate should be compared to the 15-20 percent annual growth rates of the smallest firms in the US economy estimated from the BITS data by Rossi-Hansberg and Wright, 2007. According to their estimates, the annual growth rate of firms with more than 10 employees is close to zero. Thus, we find that the customer markets feature of our model can account for around 1/3 of the growth differential between small and large firms.

As in Hopenhayn (1992), Ericson and Pakes (1995), or Luttmer (2004), idiosyncratic productivity shocks give rise to a selection effect through which small firms

[14] The density at the minimum value of s is the entry rate given our assumption that all firms enter at this value.

grow faster than mature firms. Recall that the entry decision is made prior to the draw of the entrants' initial level of productivity. Entrants that draw a low level of productivity immediately leave the market and do therefore not appear in the distribution of survivors from which employment growth rates are calculated.

The customer markets feature of our model propagates the selection effect. In particular, subject to survival, small firms grow fast because they invest in their customer stock by setting low prices. This feature is reflected in the dynamics of markups and of the customer stock that we have discussed above. This second feature is key in the sense that without it, the model is a simple extension of Hopenhayn (1992) to monopolistic competition but without the latter having any important consequences for firm dynamics. In particular, this model relies on a pure selection effect and has no source of scale dependence.

4.3 The Role of Habits and Sensitivity Analysis

The key aspect of our theory is the presence of deep habits. We will now examine the extent to which this aspect is important for our results. Moreover, as is clear from Section 3, the calibration of the model is associated with substantial uncertainty. Therefore we also investigate the importance of key structural parameters.

In Figure 10 we illustrate the firm distribution and the relationship between firm size and firm growth when we set θ equal to -0.05 which implies a very limited role of deep habits. For such values of θ firms set markups that are basically independent of marginal costs and the size of the customer stock. In this case, the firm size distribution is very concentrated, basically flat and bears little or no resemblance to a Pareto distribution. Panel B illustrates the conditional growth rate and we see that, while there is some minor role of selection, there is little difference in the implied growth rates of small and big firms. This result is partly due to our calibration of the stochastic process for the idiosyncratic productivity shock and to the size of entry and exit costs. More precisely, more persistent and volatile productivity shocks gives rise to a stronger selection effect which is reinforced by larger operational costs. However, we find it interesting that the deep habits mechanism propagates the selection effect very strongly even in the presence of relatively modest persistence and volatility of the productivity shocks. In short, they key aspect of our results is the allowance for customer markets effects through deep habits.

The key uncertainty of our calibration is related to the idiosyncratic productivity shock process. As we discussed in Section 3, estimates of the persistence of technology shocks and of the volatility of the innovation term are associated with substantial amount of uncertainty. In Figures 11 and 12 we examine the sensitivity of the firm size distributions and of the conditional employment growth rates to the volatility of the innovations to this process and to the persistence of the process. In Figure 11 we half the standard deviation of the volatility of the process setting $\sigma = 0.02$ and in Figure 12 we lower the persistence of process to 0.7. Both these calibrations lower the overall volatility of TFP across firms with 50 percent.

The inequality in the firm size distribution remains for either of the alternative calibrations of the technology shock process. However, lowering the volatility of the technology shock process by either lowering its persistence or the volatility of the innovations implies less scale dependence in firm employment growth rates. The intuition for this result is slightly different for the two cases. When the persistence of the shock is low, regardless of the initial draw of productivity, there is a substantial probability that next periods productivity is high. This implies less selection and less propagation of selection through customer stocks since small firms that experience a drop in their productivity might find it optimal to remain in the market. When the volatility of the innovations is low, the implied differences in the values of TFP across firms is also low. Therefore, only the firms with the very lowest levels of productivity exit the market leaving little room for selection.

These results indicate that it will be important to obtain more precise estimates of the idiosyncratic technology shock process. However, this sensitivity to this aspect of the calibration is common to many models of firm dynamics that rely on selection and is therefore not special to our model.

4.4 Implications for the Cross-Section of Returns

To be added

5 Conclusions and Summary

This paper has suggested that imperfect competition structures and customer markets effects may help to understand the unequal size distribution of firms and the scale dependence in firm growth rates. We studied a model with deep habits in which consumers form habits to individual goods. The deep habits mechanism has been shown to be successful for accounting for the effects of aggregate demand shocks in settings without firm heterogeneity. Key to this theory is that the presence of habits in consumers' preferences implies dynamic pricing models on the part of firms. Firms, knowing that their current prices affect future demand for their products, set markups that vary in response to changes in marginal costs and aggregate demand.

In this paper we have introduced this aspect into a model of firm turnover and we have shown that it has interesting implications. We studied an environment in which firms enter the market with small customer stocks and in which firms are subject to idiosyncratic productivity shocks over their life-cycle. Incumbent firms may exit the market when their continuation value falls below their outside option and new entrants will enter as long as the expected value of entry is positive. The existence of fixed operational costs and fixed entry costs implies that there is firm turnover in the face of idiosyncratic productivity shocks. We studied the properties of the stationary equilibrium in which the firm distribution is invariant. The model was calibrated and we investigated its quantitative implications. In the calibrated economy, entrants with high productivity set low markups in order build up their customer stock. Large customer stocks help firms insure themselves against adverse productivity shocks. As the firm grows, it exploits its large customer stock and increases its markup.

Therefore, the customer markets feature of our model propagates the selection effects of idiosyncratic productivity shocks. This mechanism is quantitatively relevant. We showed that it helps significantly account for the fact that small firms grow faster than big firms and for the inequality in the firm size distribution. In particular, we showed that this feature by itself may account for at least one third of the 15 percent difference in employment growth rates of small and big firms that is observed in U.S. data. Naturally, there are many other empirically relevant mechanisms that might lead to scale dependent firm dynamics, but we believe that the habit formation model studied in this paper may be an interesting candidate.

The model has also potentially interesting implications for other issues. In particular, we believe that it is worthwhile to explore its implications for the relationship between firm values and firm returns and for the adjustment of prices. Due to the dynamics of markups and customer stocks the theory has non-trivial consequences for the relationship between firm value and profits. Short but productive firms may sacrifice current profits for future profits and therefore produce low current returns while still having high present discounted values of future returns. In contrast, firms with little hope of survival may sacrifice future profits for current profits but their value is low. Therefore, the model has potentially much more interesting asset pricing implications than standard monopolistic competition models in which constant markup policies lead to monotone relationships between firm values and current profits. The dynamics of markups may also have interesting ramifications on the distribution of price changes. We will explore these issues in future research.

References

[1] Albuquerque, Rui, and Hugo A. Hopenhayn, 2004, “Optimal Lending Contracts and Firm Dynamics”, *Review of Economic Studies 71*, 285-315.

[2] Clementi, Gian Luca, and Hugo A. Hopenhayn, 2006, “A Theory of Financing Constraints and Firm Dynamics”, *Quarterly Journal of Economics 121*, 229-65.

[3] Comin, Diego, and Thomas Philippon, 2005, “The Rise in Firm Level Volatility: Causes and Consequences”, NBER working paper no. 11388.

[4] Cooley, Thomas F., and Vincenzo Quadrini, 2001, “Financial Markets and Firm Dynamics”, *American Economic Review 91*, 1286-1310.

[5] Cooper, Russell W., John C. Haltiwanger, and Jonathan Willis, 2004, “Dynamics of Labor Demand: Evidence from Plant Level Observations and Aggregate Implications”, NBER working paper no. 10297.

[6] Davis, Steven J., John Haltiwanger, and Ron Jarmin, 2006, “Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms”, CES working paper no. 06-17.

[7] Ericson, Richard, and Ariel Pakes, 1995, “Markov-Perfect Industry Dynamics: A Framework for Empirical Work”, *Review of Economic Studies 62*, 53-82.

[8] Garicano, Luis, and Esteban Rossi-Hansberg, 2004, “Inequality and the Organization of Knowledge”, *American Economic Review P&P 94*, 197-202.

[9] Hopenhayn, Hugo A., 1992, “Entry, Exit and Firm Dynamics in the Long Run”, *Econometrica 60*, 1127-1150.

[10] Jaimovich, Nir, 2004, “Firm Dynamics, Markup Variation, and the Business Cycle”, working paper, UC San Diego.

[11] Jovanovic, Boyan, 1982, “Selection and the Evolution of Industry”, *Econometrica 50*, 649-70.

[12] Khan, Aubhik, and Julia Thomas, 2007, “Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics”, NBER working paper no. 12845.

[13] Klemperer, Paul, 1995, “Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics and International Trade”, *Review of Economic Studies 62*, 515-39.

[14] Klette, Tor Jakob, and Sam Kortum, 2004, “Innovating Firms and Aggregate Innovations”, *Journal of Political Economy 112*, 986-1018.

[15] Lucas, Robert E. Jr., 1978, "On the Size Distribution of Business Firms", *Bell Journal of Economics 9*, 508-23.

[16] Luttmer, Erzo, 2007, "Selection, Growth and the Size Distribution of Firms", forthcoming, *Quarterly Journal of Economics.*

[17] Martins, J., S. Scapetta and D. Pilat, 1996, "Markup Pricing, Market Structure and the Business Cycle", *OECD Economic Studies 27*, 71-105.

[18] Nakamura, Emi, and Jon Steinsson, 2006, "Price Setting in Forward-Looking Customer Markets", mimeo, Harvard University.

[19] Phelps, Edmund S. and Sidney G. Winter, 1970, "Optimal Price Policy under Atomistic Competition", in Edmund S. Phelps, Ed., *Microeconomic Foundations of Employment and Inflation Theory*, New York, NY: W.W. Norton.

[20] Ravn, Morten O., Stephanie Schmitt-Grohe and Martin Uribe, 2006, "Deep Habits", *Review of Economic Studies 73*, 195-218.

[21] Ravn, Morten O., Stephanie Schmitt-Grohe and Martin Uribe, 2007, "Incomplete Cost Pass-Through under Deep Habits", NBER working paper no. 12961.

[22] Rossi-Hansberg, Esteban, and Mark L.J. Wright, 2007, "Establishment Size Dynamics in the Aggregate Economy", forthcoming, *American Economic Review.*

[23] Sutton, John, 1997, "Gibrat's Legacy", *Journal of Economic Literature 35*, 40-59.

[24] Tauchen, George, 1986, "Finite State Markov Chain Approximations to Univariate and Vector Autoregressions", *Economics Letters 20*, 177-81.

[25] World Bank, 2007, "Doing Business", on-line version.

6 Appendix 1: Derivation of Equation (16)

We simplify the model by neglecting the exit decision. Bellman's equation for an incumbent firm can be expressed as:

$$V(s_i, A_i) = \max_{p_i, c_i, h_i, s_i'} \left[\frac{p_i}{P} c_i - w h_i - \varphi + \beta E V(s_i', A_i') \right]$$

subject to:

$$\begin{aligned} s_i' &= (1-\rho) s_i + \rho c_i \\ c_i &= \left(\frac{p_i}{P} \right)^{-\eta} s_i^{\theta(1-\eta)} X \\ c_i &= A_i h_i^{\alpha} \end{aligned}$$

The first-order conditions for c_i, p_i, h_i, s_i' and the envelope condition (in that order) are:

$$\begin{aligned} c_i &: \frac{p_i}{P} - \lambda_c - \lambda_y + \rho \lambda_s = 0 \\ p_i &: \lambda_c = \frac{1}{\eta} \frac{p_i}{P} \\ h_i &: \lambda_y = \frac{w}{\alpha A_i h^{\alpha-1}} \\ s_i' &: \beta E \frac{\partial V'}{\partial s_i'} = \lambda_s \\ s &: V_s = \lambda_s (1-\rho) + \theta (1-\eta) \lambda_c \frac{c_i}{s_i} \end{aligned}$$

where λ_s is the multiplier on the transition rule for s_i, λ_c is the multiplier on the firm's demand function, and λ_y is the multiplier on the production function. The latter has the interpretation of the firm's marginal costs. Using the first-order condition for p_i in the first-order condition for c_i gives us:

$$\lambda_s = \frac{1}{\rho} \left(\lambda_y + \left(\frac{\eta - 1}{\eta} \right) \frac{p_i}{P} \right) \tag{20}$$

Next, combining the condition for s_i' with the envelope condition gives:

$$\lambda_s = E \beta \left[\lambda_s' (1-\rho) + \theta (1-\eta) \frac{1}{\eta} \frac{p_i'}{P'} \frac{c_i'}{s_i'} \right]$$

Iterating this expression forwards (using date notation) and imposing a transversality condition implies that:

$$\lambda_{s,t} = \beta E_t \left[\theta (1-\eta) \sum_{j=1}^{\infty} (\beta (1-\rho))^{j-1} \lambda_{c,t+j} \frac{c_{i,t+j}}{s_{i,t+j}} \right]$$

Finally, substituting the expression for λ_s from equation (20) into this expression and re-arranging gives equation (16) in the text.

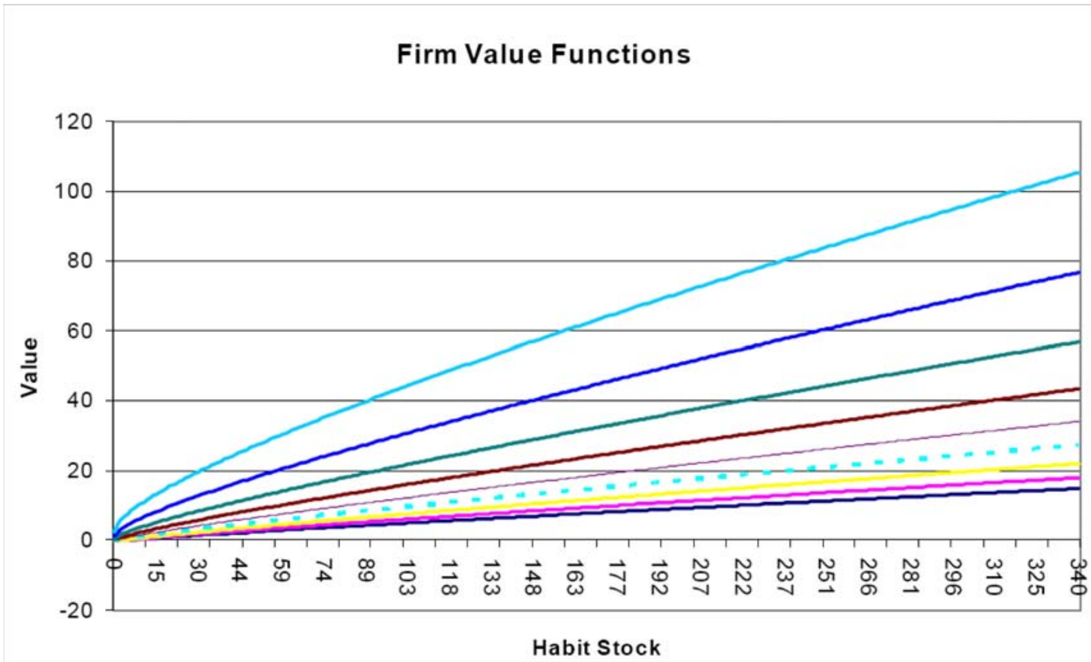

Figure 1: The value of the firm contingent upon its level of productivity. The higher the value function, the higher the level of productivity.

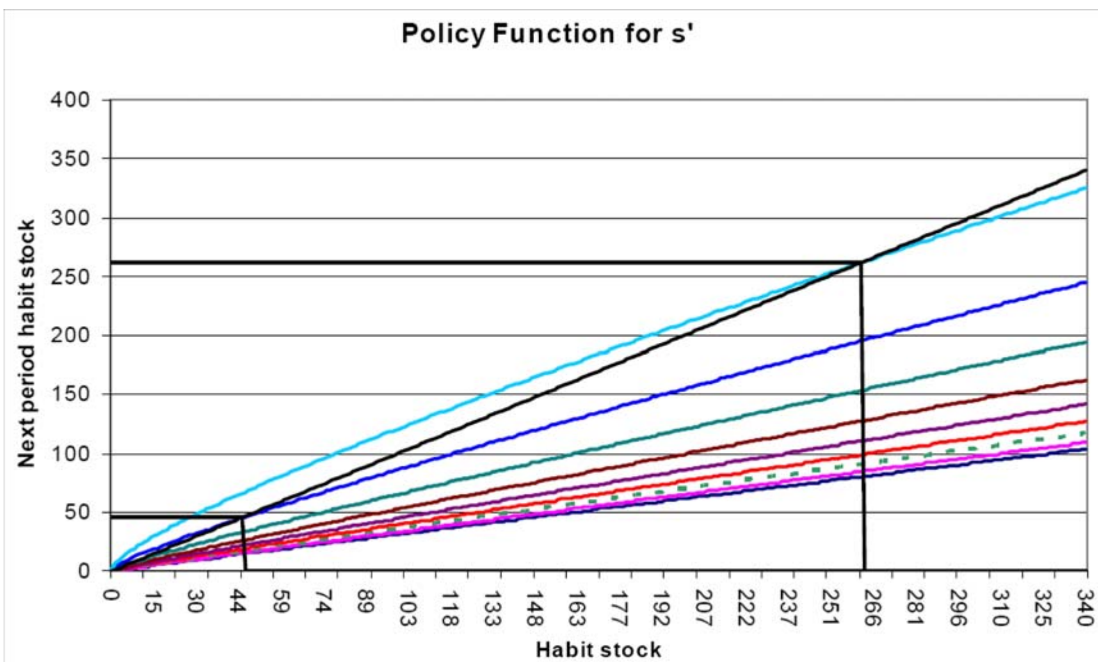

Figure 2: The firms' policy function for $s'(s, A)$. Higher levels of productivity are associated with higher policy functions.

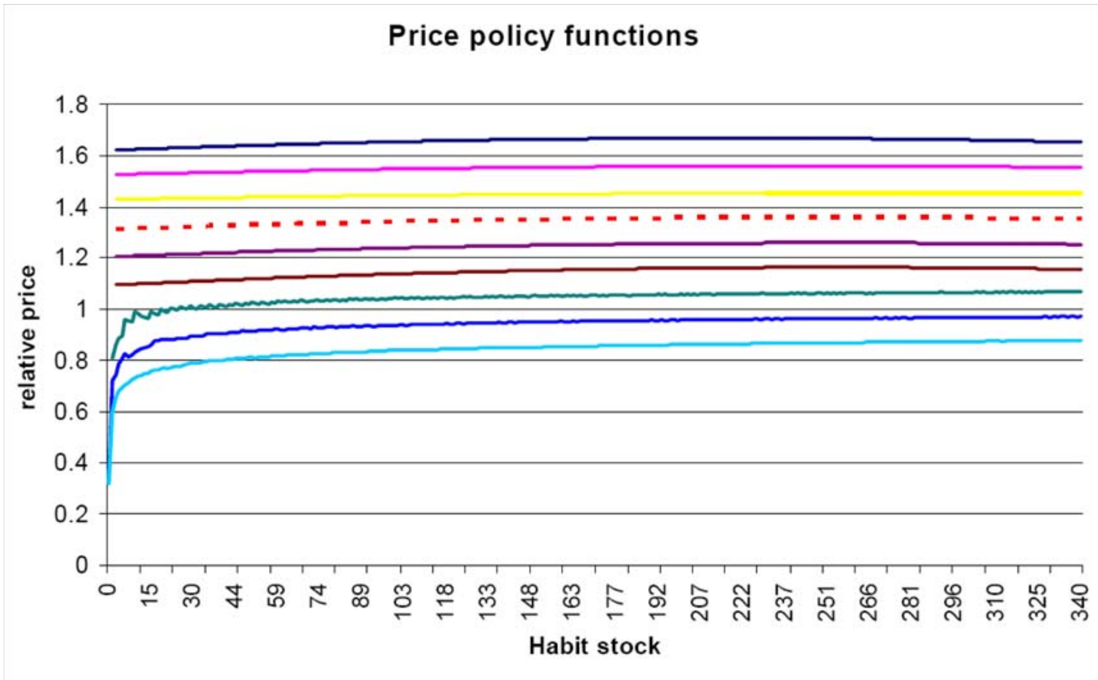

Figure 3: Policy function for prices. Higher levels of productivity are associated with lower policy functions.

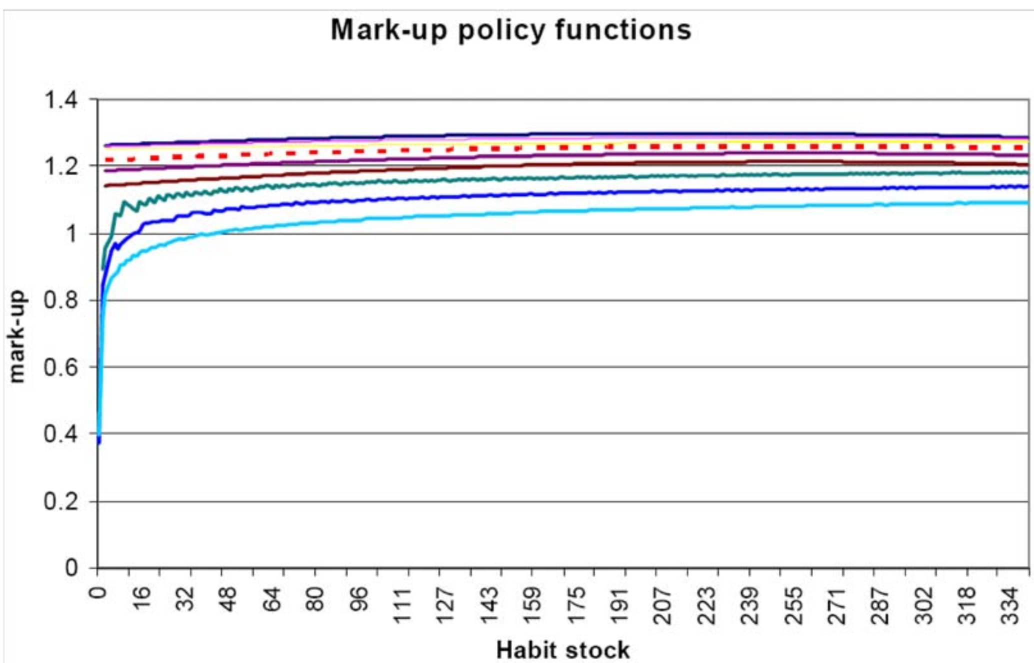

Figure 4: Policy function for mark-ups. Higher levels of productivity are associated with lower policy functions.

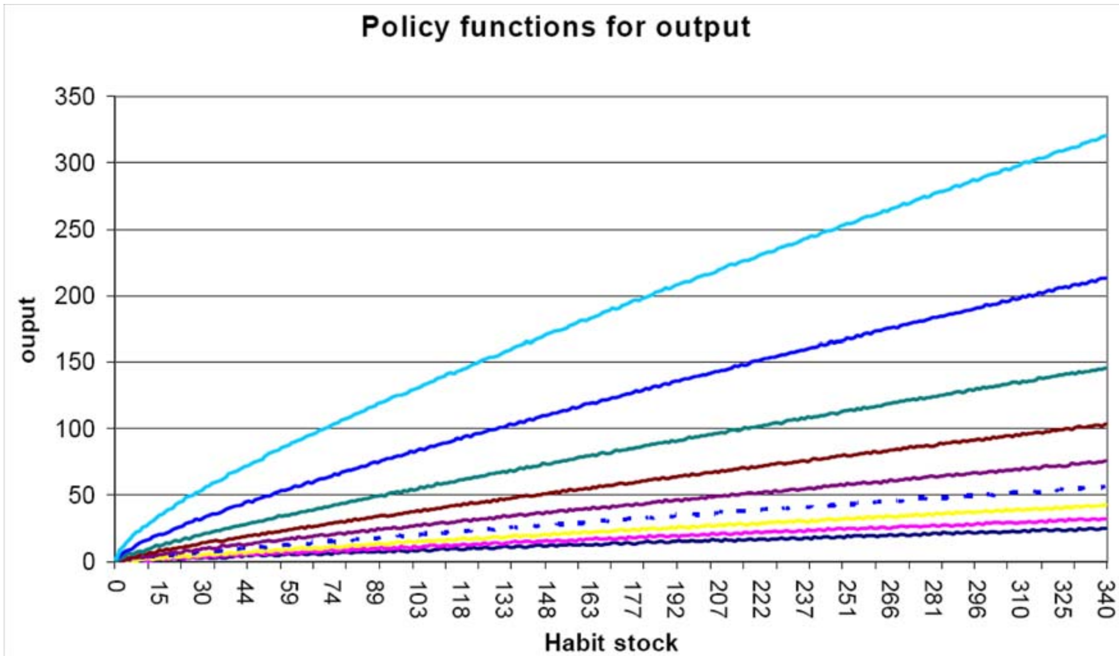

Figure 5: Policy function for output. Higher levels of productivity are associated with higher policy functions.

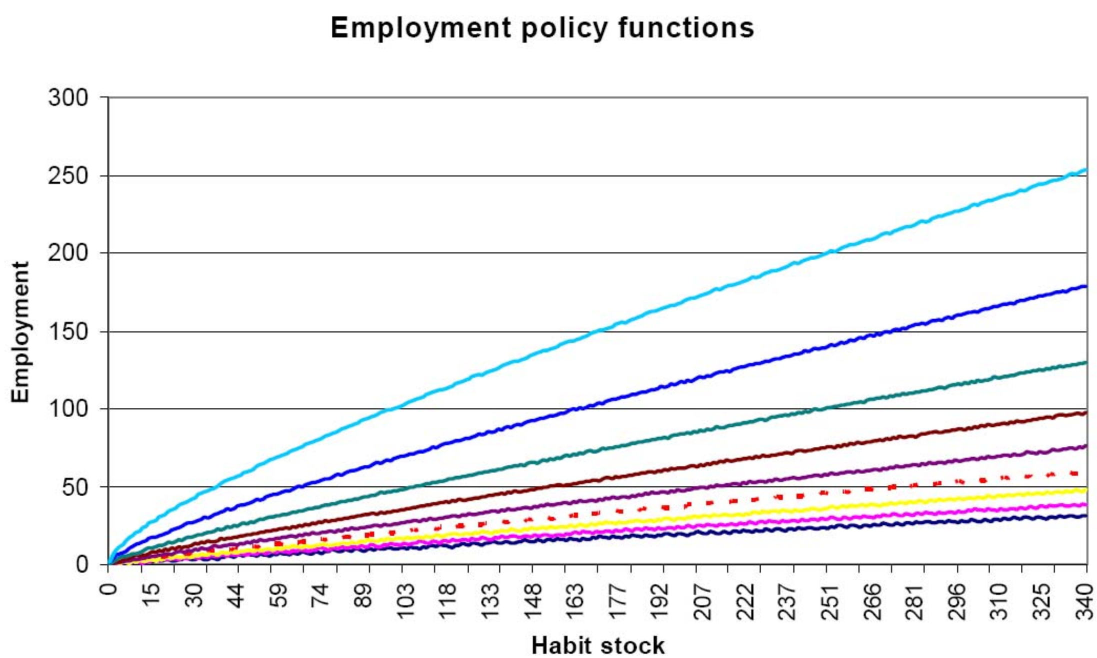

Figure 6: Policy function for employment. Higher levels of productivity are associated with higher policy functions.

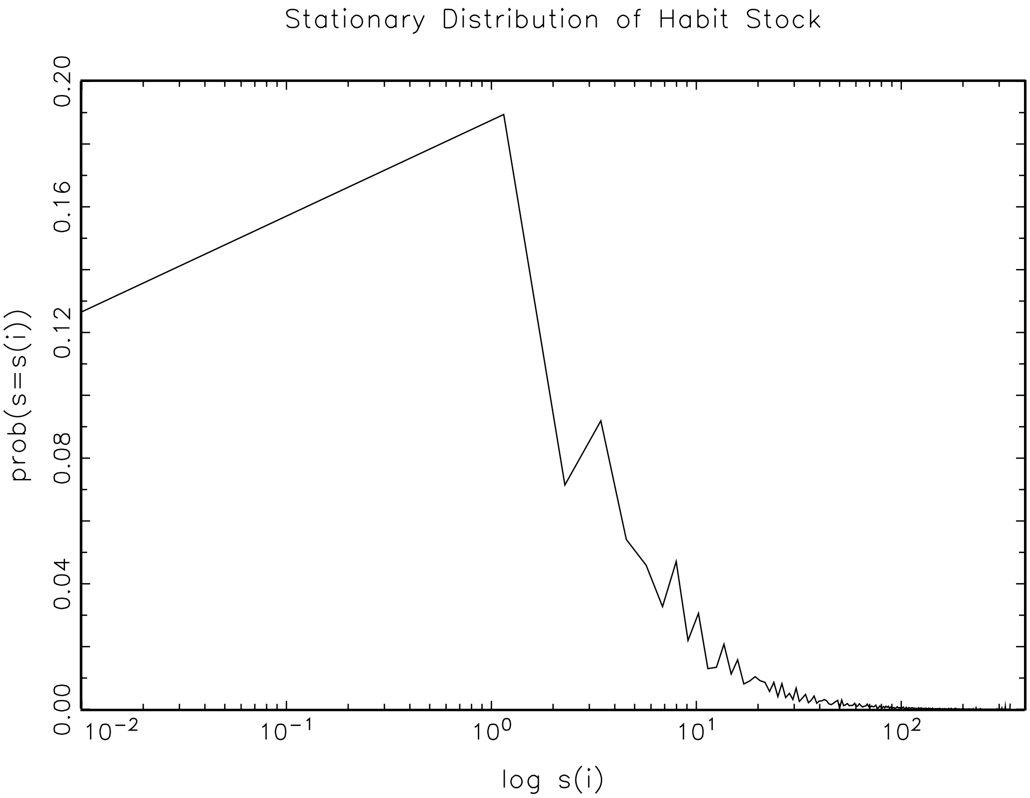

Figure 7: The stationary distribution of the habit stock.

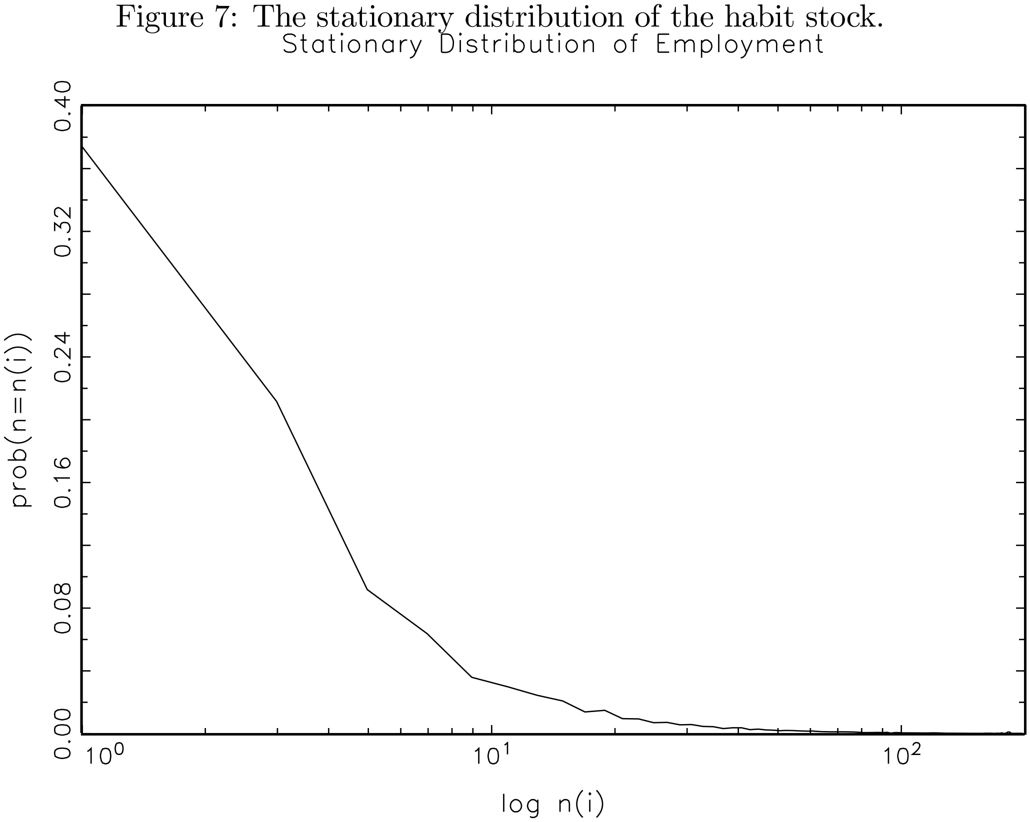

Figure 8: The stationary distribution of employment.

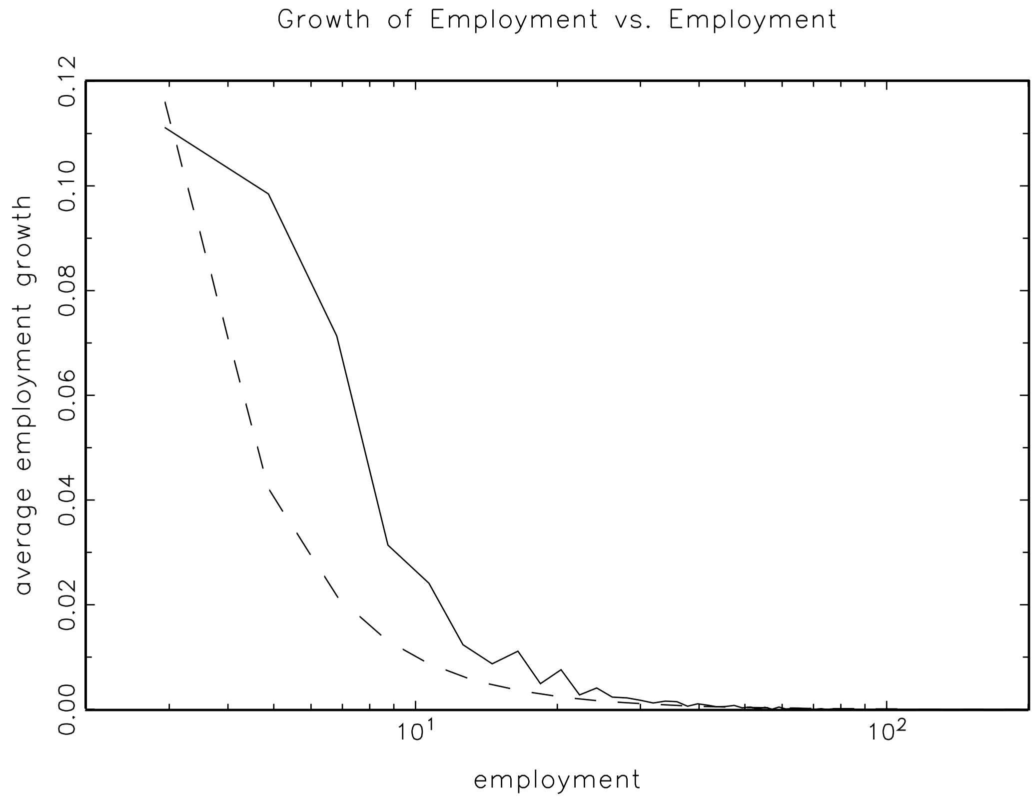

Figure 8A: The firm size distribution and the Pareto distribution with parameter 1 (the dashed line).

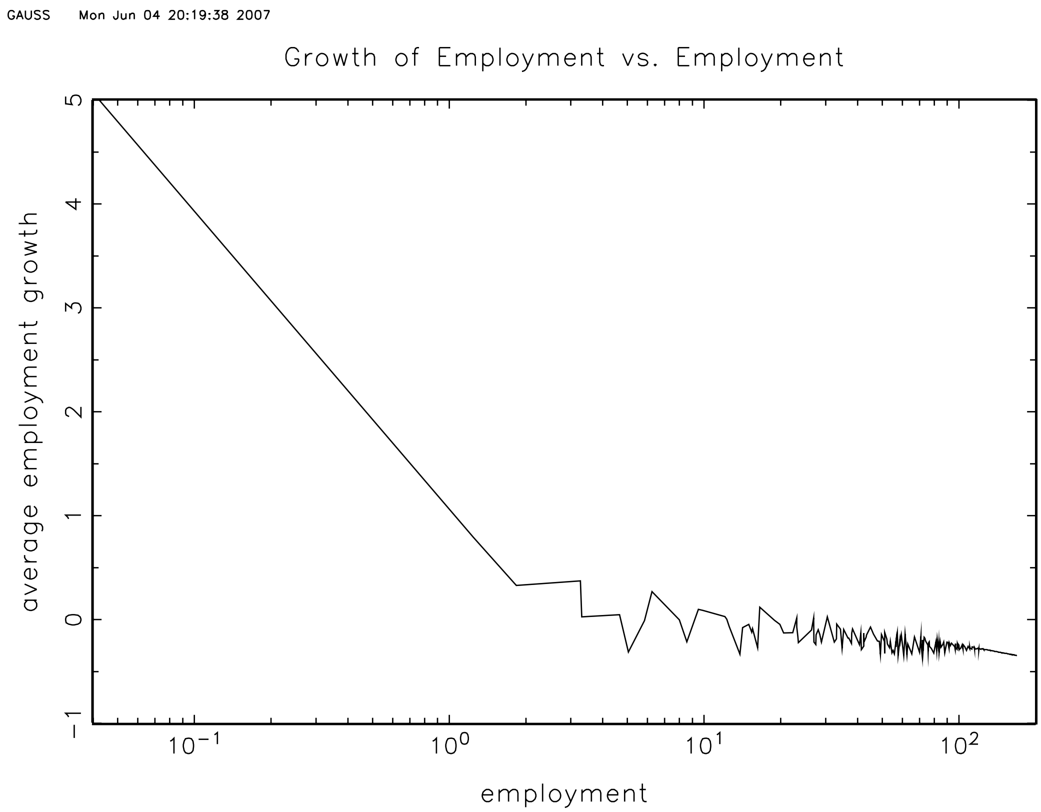

Figure 9: The conditional growth rate of employment.

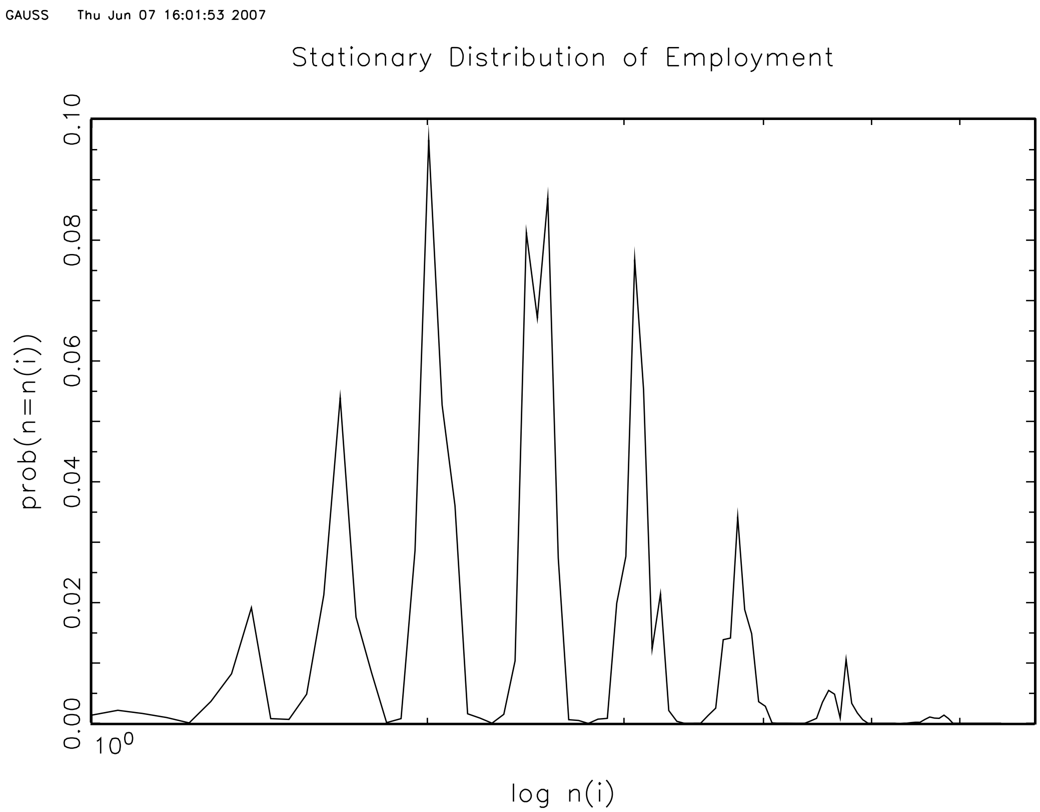

Figure 10a: The firm size distribution when $\theta = -0.05$

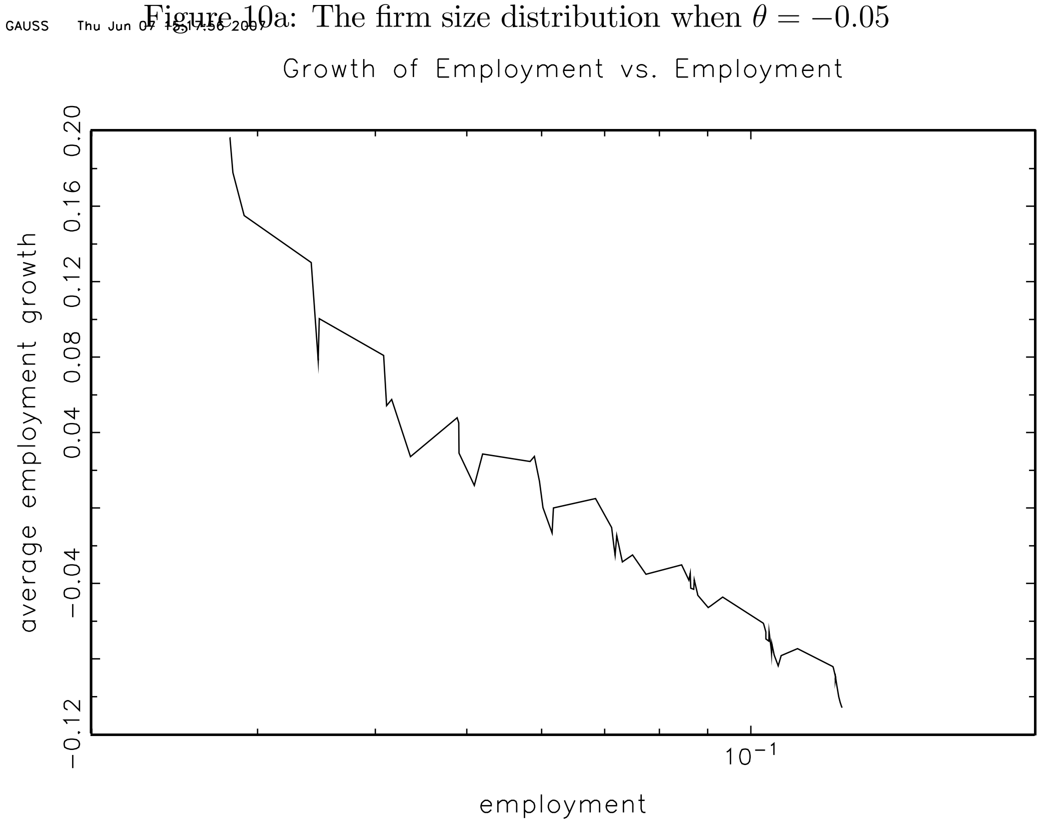

Figure 10b: The conditional growth rate of employment when $\theta = -0.05$

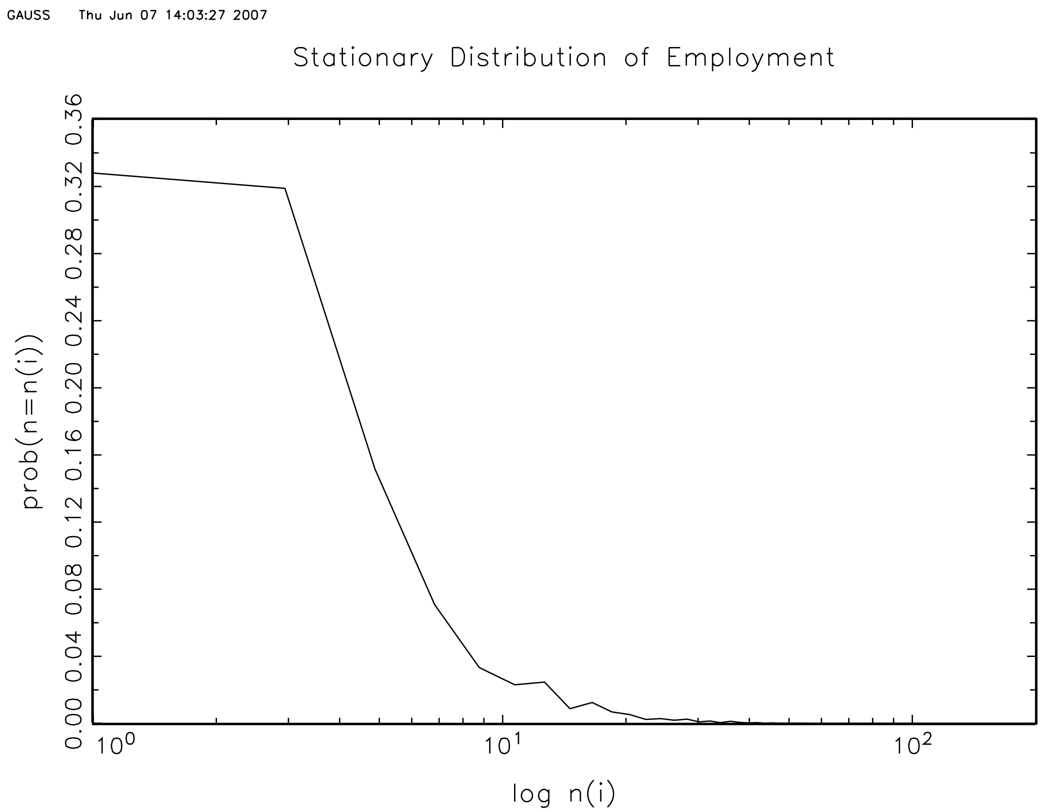

Figure 11a: The firm size distribution with low volatility

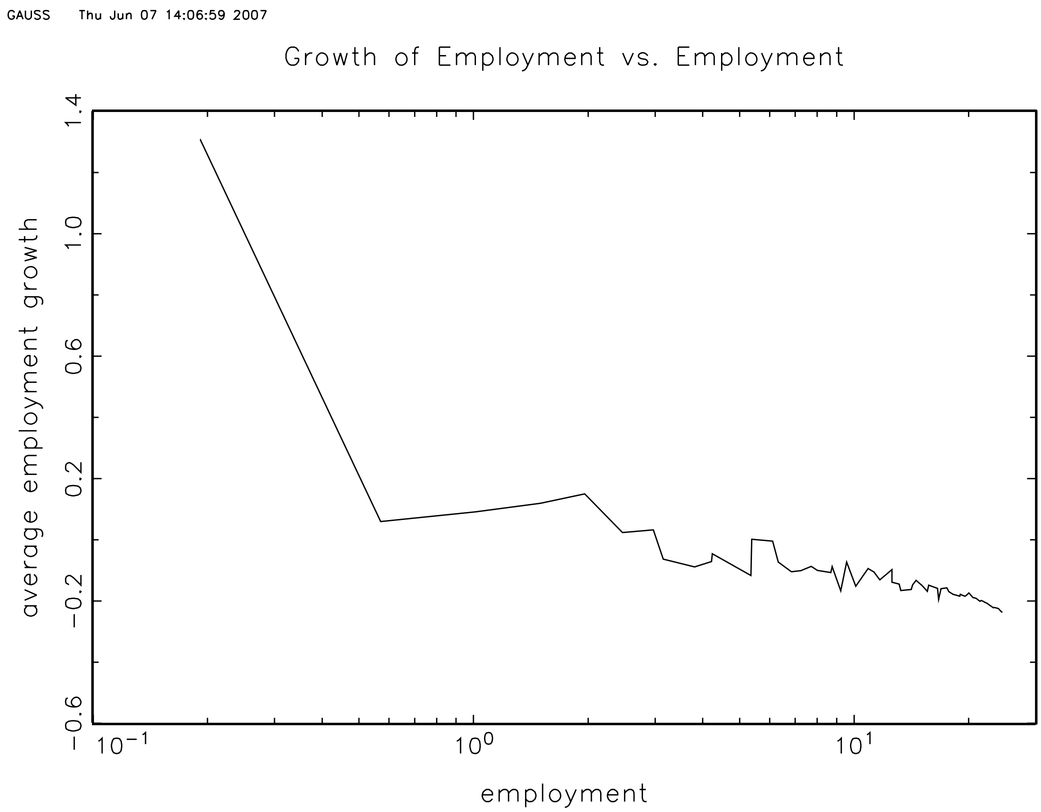

Figure 11b: The conditional growth rate of employment with low volatility

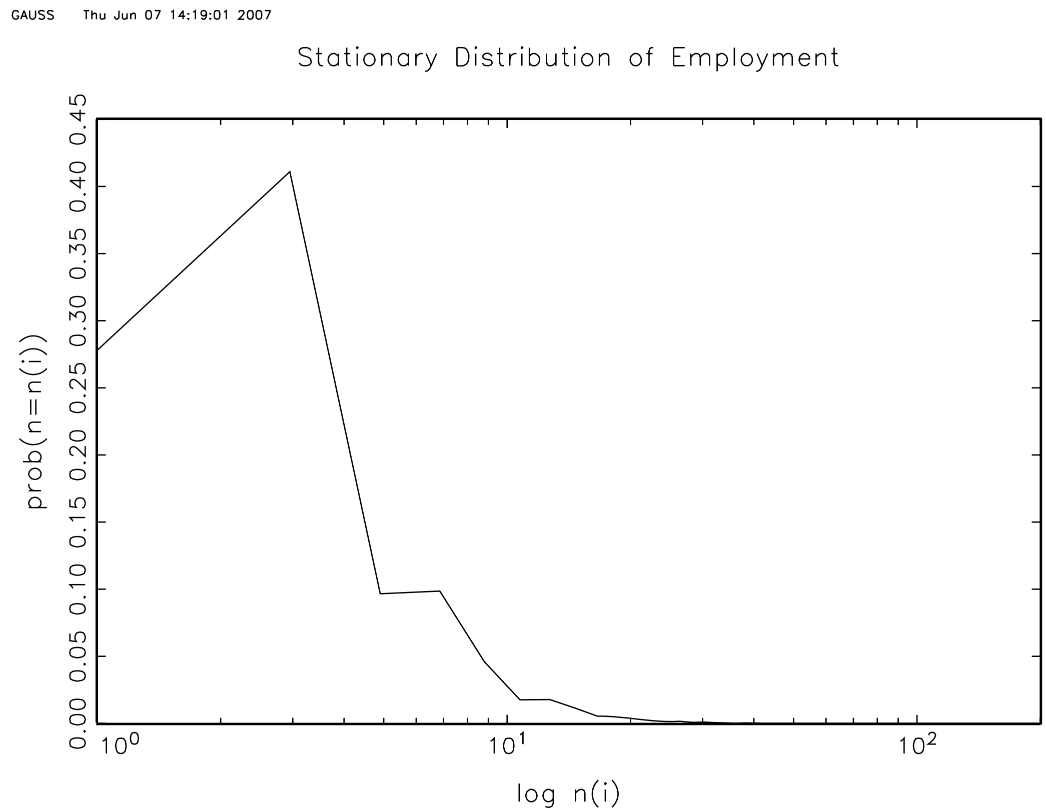

Figure 12a: The firm size distribution when technology shocks have low persistence

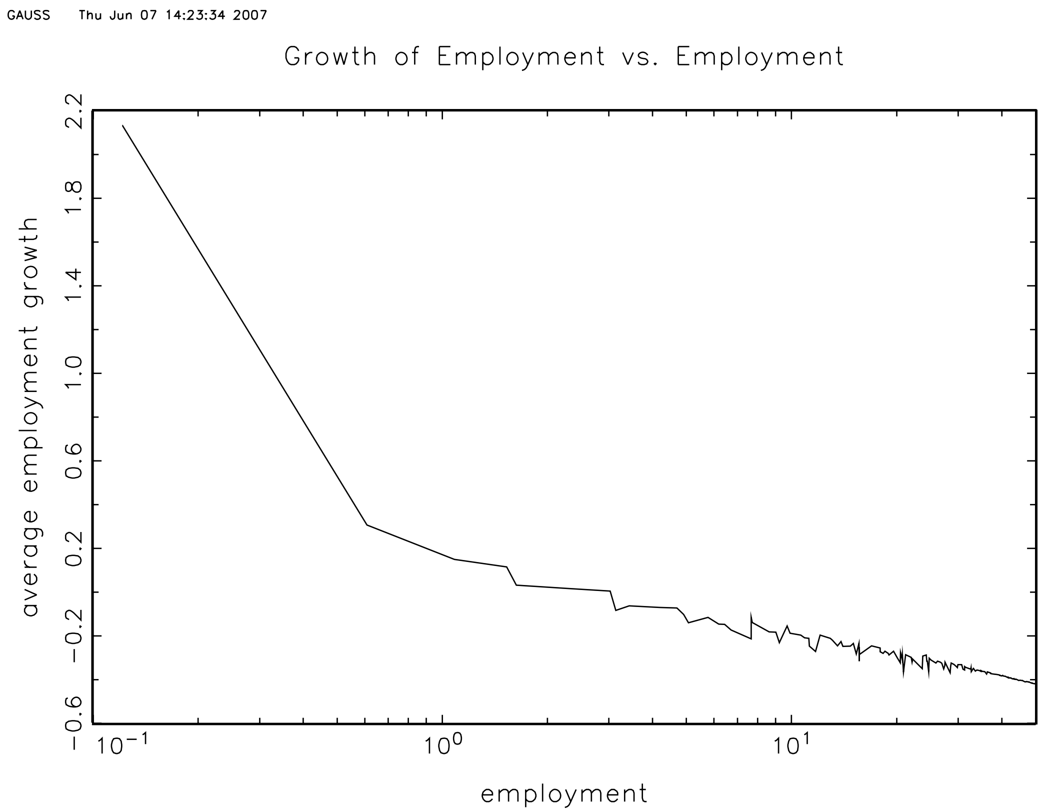

Figure 12b: The conditional growth rate of employment with low persistence