Firm Dynamics with Lumpy Adjustment and Learning

Eugénio Pinto*
University of Maryland

First Version: April 2004
This Version: February 2005

PRELIMINARY

Abstract

In a model of Bayesian learning about efficiency, we consider the effect of entry, regular, and exit nonconvex adjustment costs, both in the form of proportional and fixed costs. We conclude that, except for entry and hiring fixed adjustment costs, firms will have a tendency to start smaller, to grow faster, and, in some circumstances, to display a smaller exit propensity. The existence of a distribution of productivities highly skewed to the right, and of a technology with high (but decreasing) returns to labor will intensify the above predictions. Therefore, two paradigms can occur. When entry and hiring fixed adjustment costs are high, most growth is due to the natural selection mechanism. When entry and hiring fixed costs are dominated by the other adjustment costs, the increase in average size is mostly due to the upward adjustment of (efficient) surviving firms. Simulations, in finite horizon, are presented.

JEL Classification: E24, L11, L16
Keywords: Adjustment Costs, Learning, Young Firms

*University of Maryland, Department of Economics, Tydings Hall, Room 3105, College Park, MD 20742 (e-mail:pinto@econ.bsos.umd.edu). I would like to thank John Shea, John Haltiwanger, and Michael Pries, for their encouragement of this project.
1 Introduction

In this paper, we use nonconvex adjustment costs and a Bayesian learning process to interpret and, eventually, explain some well known differences in dynamic behavior of young and old firms. The literature presenting evidence on young versus old firms is wide. It ranges from studies with more emphasis in macroeconomics, such as the input and output reallocation over the business cycle, to studies which focus more on industrial organization topics, such as the determinants of firms life duration. Here, we are interested in the part that deals with the differences in growth and input reallocation patterns (e.g. Timothy Dunne, Mark Roberts, and Larry Samuelson 1989a, 1989b, José Mata and Pedro Portugal 1994, José Mata, Pedro Portugal, and Paulo Guimarães 1995, Luís Cabral and José Mata 2003, Eric Bartelsman, Stefano Scarpetta, and Fabiano Schivardi 2003). In particular, we argue that nonconvex adjustment costs, in a learning-about-efficiency environment, can be an alternative to theories based on selection of efficient firms and on credit constraints, in explaining why young firms, when compared to old firms, are smaller, grow faster, and display different patterns of input reallocation.

We start by showing some age related facts for the Portuguese manufacturing sector. We use Quadros de Pessoal (QP), a micro level database containing information on almost all Portuguese firms, establishments, and workers from 1985 to 2000. This exercise reveals that there is wide dispersion in entry and exit size, and that the firm size distribution is stochastically increasing with age, mainly because survivors increase significantly in size. We also find that young and old firms exhibit different patterns of job reallocation. Young firms tend to have higher rates and account for a substantial fraction of gross job flows. They also tend to display procyclical reallocation rates and countercyclical excess reallocation rates, whereas the opposite occurs for old firms.

We then build a model that deals with the growth pattern of young firms. Even though this model assumes a Bayesian learning process, which, through a selection mechanism, as in Boyan Jovanovic (1982), can lead to higher growth among young firms, we are able to provide an alternative explanation for this fact by using nonconvex adjustment costs. We assume that entering firms are uncertain about their true efficiency and, as time goes by, use a Bayesian updating rule to gradually ascertain about it. As they update their posterior, and become more confident about their perceptions, they change their optimal decisions concerning the permanence in the industry and the input and output levels (conditional on staying). On top of this learning process, we consider the general case of entry, regular, and exit adjustment costs, and analyze the isolated contribution of each of them. The conclusion is that adjustment costs generate nontrivial effects in the dynamics of the model. To our knowledge, this is the first time that adjustment costs are considered to affect both the size distribution, and the age related growth patterns of firms.

In general, if adjustment costs are high, firms will tend to enter the industry with a smaller size, will tend to grow faster, and will tend to display smaller exit rates. The idea behind these results is the following. The initial uncertainty about true profitability makes entering firms prudent, that is, they enter smaller and “wait and see”, since they want to avoid future overcapacity. Future overcapacity means that they would have incurred superfluous initial entering/hiring costs and firing/shutdown costs on all excess inputs, namely workers. As
they tend to enter smaller, they tend to grow faster, even though there is less firms exiting the market, and therefore less selection of inefficient firms.

Two aspects of the problem that intensify the above mentioned tendency are a distribution of productivity highly skewed to the right, and a production technology with high (but decreasing) returns to labor. Firms know that most certainly their productivity will not be very high, and that they will be more propense to adjust upwards. Therefore, adjustment costs and uncertainty about the true efficiency in the initial years of activity makes them prudent, so that they will start smaller, but will gradually catch up as the uncertainty is resolved as they get older.

In a discrete time and infinite horizon model, we derive the Euler equations for optimal employment adjustments, and show the effect of entry, regular and exit adjustment costs on the optimal employment policy over time. Two possible paradigms can occur. The first corresponds to a situation where entry and hiring fixed costs are relatively significant. In this case, most growth will be due to the natural selection mechanism. The second corresponds to a situation where entry and hiring fixed costs are dominated by the other, proportional and/or fixed, adjustment costs. In this case, the more important part of growth will be due to growth of surviving firms. To illustrate these assertions, we do some simulations of a finite learning horizon version of the model, assuming a lognormal distribution of the efficiency parameter and a Cobb-Douglas production function.

This makes our argument a potential challenge to theories based on financing constrains. In the Portuguese manufacturing sector most growth of young firms is accounted for by the growth of survivors and not to by the exit of less efficient and small firms. Therefore, the question remains, if in circumstances where most growth is due to growth of survivors, if that results from financing constrains of young firms, or if, instead, that is the result of young firms being "prudent" because they face adjustment costs in a learning environment. That is, young firms have an incentive to start small because by doing that they are balancing lost profits with the expected decrease in adjustment costs.

Some evidence for this phenomenon is provided by Mata et al. (1995, p. 478):

One possible explanation for [the result that after controlling for current size, initial size, and therefore past growth, is important] is that current size is likely to be different from desired size .... This divergence may arise if, for example, there are adjustment costs that prevent firms from fully adjusting their size instantaneously in response to an observed market signal. ... Therefore, our results may be interpreted as suggesting the existence of adjustment costs in the process of post-entry growth.

We organize our work in five sections and one appendix. In section 2, we present some of the stylized facts concerning the differences in dynamic behavior between young and old firms, and some of the theoretical models that have been put forward to explain them. In section 3, we present a modification of the Jovanovic (1982) model of industry dynamics by incorporating entry, regular and exit adjustment costs. In section 4, we perform a simulation of the model by assuming a finite learning horizon a lognormal distribution for the efficiency parameter and a Cobb-Douglas production function. In section 5, we leave a brief summary of the current empirical projects related to this paper. In the conclusion, we present our main
Table 1: 1988 Firm Cohort: Current Year Size Distribution of Survivors

<table>
<thead>
<tr>
<th>Year</th>
<th>n</th>
<th>mean</th>
<th>sd</th>
<th>sk</th>
<th>p5</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p95</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>4834</td>
<td>10.3</td>
<td>30.0</td>
<td>14.5</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>32</td>
</tr>
<tr>
<td>1989</td>
<td>3779</td>
<td>12.7</td>
<td>32.3</td>
<td>11.4</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>12</td>
<td>40</td>
</tr>
<tr>
<td>1990</td>
<td>3215</td>
<td>15.0</td>
<td>37.2</td>
<td>10.5</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>14</td>
<td>50</td>
</tr>
<tr>
<td>1991</td>
<td>2887</td>
<td>15.8</td>
<td>35.9</td>
<td>9.3</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>15</td>
<td>55</td>
</tr>
<tr>
<td>1992</td>
<td>2692</td>
<td>16.3</td>
<td>38.1</td>
<td>9.3</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>15</td>
<td>61</td>
</tr>
<tr>
<td>1993</td>
<td>2475</td>
<td>15.6</td>
<td>36.6</td>
<td>10.4</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>14</td>
<td>54</td>
</tr>
<tr>
<td>1994</td>
<td>2268</td>
<td>16.3</td>
<td>43.7</td>
<td>12.3</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>14</td>
<td>54</td>
</tr>
<tr>
<td>1995</td>
<td>2167</td>
<td>16.6</td>
<td>42.7</td>
<td>11.7</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>14</td>
<td>57</td>
</tr>
<tr>
<td>1996</td>
<td>1929</td>
<td>17.1</td>
<td>41.9</td>
<td>10.9</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>16</td>
<td>61</td>
</tr>
<tr>
<td>1997</td>
<td>1950</td>
<td>17.4</td>
<td>41.8</td>
<td>10.7</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>16</td>
<td>63</td>
</tr>
<tr>
<td>1998</td>
<td>1871</td>
<td>17.7</td>
<td>43.8</td>
<td>11.7</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>15</td>
<td>65</td>
</tr>
<tr>
<td>1999</td>
<td>1812</td>
<td>18.3</td>
<td>45.3</td>
<td>11.7</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>16</td>
<td>67</td>
</tr>
</tbody>
</table>

Notes: *n* is the number of survivors (without a missing number of employees); *mean*, *sd*, *sk*, and *px* are the mean, standard deviation, skewness coefficient and *x*-th percentile of the survivors size distribution, respectively.

achievements, and projects for future work. We also have an appendix with three sections. Appendix A deals with the proof of propositions and corollaries in the main text. Appendix B contains an example of one of the arguments stated in section 3. Finally, appendix C contains some details concerning the computational algorithm used to simulate the finite learning horizon version of the model.

## 2 Firm Behavior Across Age Cohorts

### 2.1 Some Facts

There is a well established literature both on the identification and the explanation of differences in the average characteristics displayed by young and old firms. In this subsection, we present and review some of the stylized facts in question, and, in the next subsection, we will comment on some of the theories that attempt to explain them.

In tables 1 to 3, we present some characteristics of the size distribution for a cohort of entering firms. The data is taken from *Quadros de Pessoal*, a Portuguese database, originating from a mandatory annual survey run by the Ministry of Employment, containing information on almost all firms with paid employees, and covering the period from 1985 until 2000.¹

From table 1, we can see that the current year size distribution of survivors is highly skewed to the right for all ages, and that it stochastically increases with age: the mean increases from 10.3 in 1988 to 18.3 in 1999, the median increases from 4 to 7, the standard deviation increases significantly, the 95-th percentile more than doubles, and the decrease in

---

¹We have selected the 1988 entering cohort in order to better detect false entries, and are only displaying results up to 1999 so that most false exits are avoided.
Table 2: 1988 Firm Cohort: 1988 Size Distribution of Survivors

<table>
<thead>
<tr>
<th>Year</th>
<th>n</th>
<th>er</th>
<th>ccer</th>
<th>mean</th>
<th>sd</th>
<th>sk</th>
<th>p5</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p95</th>
<th>p95</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>4130</td>
<td>14.6</td>
<td>14.6</td>
<td>10.2</td>
<td>27.2</td>
<td>14.6</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>3672</td>
<td>11.1</td>
<td>24.0</td>
<td>10.5</td>
<td>28.4</td>
<td>14.3</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>3350</td>
<td>8.8</td>
<td>30.7</td>
<td>10.6</td>
<td>25.7</td>
<td>10.4</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>3098</td>
<td>7.5</td>
<td>35.9</td>
<td>10.7</td>
<td>26.4</td>
<td>10.3</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>2855</td>
<td>7.8</td>
<td>40.9</td>
<td>10.3</td>
<td>25.6</td>
<td>10.9</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>2592</td>
<td>9.2</td>
<td>46.4</td>
<td>10.3</td>
<td>25.7</td>
<td>11.2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>2443</td>
<td>5.7</td>
<td>49.5</td>
<td>10.4</td>
<td>26.1</td>
<td>11.2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>2293</td>
<td>6.1</td>
<td>52.6</td>
<td>10.6</td>
<td>26.8</td>
<td>10.9</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>2176</td>
<td>5.1</td>
<td>55.0</td>
<td>10.8</td>
<td>27.4</td>
<td>10.8</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>2045</td>
<td>6.0</td>
<td>57.7</td>
<td>10.9</td>
<td>27.8</td>
<td>10.8</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>1914</td>
<td>6.4</td>
<td>60.4</td>
<td>11.0</td>
<td>28.5</td>
<td>10.6</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>

Notes: n is the number of survivors; er, and ccer are the current year and cumulative exit rates, respectively; mean, sd, sk, and px are the mean, standard deviation, skewness coefficient and x-th percentile of the survivors 1988 size distribution, respectively.

Table 3: 1988 Firm Cohort: Previous Year Size Distribution of Non Survivors

<table>
<thead>
<tr>
<th>Year</th>
<th>n</th>
<th>mean</th>
<th>sd</th>
<th>sk</th>
<th>p5</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p95</th>
<th>p95</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>704</td>
<td>11.0</td>
<td>42.8</td>
<td>12.1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>458</td>
<td>8.4</td>
<td>13.6</td>
<td>5.8</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>322</td>
<td>10.5</td>
<td>41.2</td>
<td>16.0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>252</td>
<td>11.6</td>
<td>16.8</td>
<td>3.4</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>14</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>243</td>
<td>16.9</td>
<td>38.2</td>
<td>7.6</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>18</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>263</td>
<td>11.0</td>
<td>20.2</td>
<td>7.3</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>13</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>149</td>
<td>7.9</td>
<td>12.7</td>
<td>4.0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>150</td>
<td>7.1</td>
<td>8.9</td>
<td>2.8</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>117</td>
<td>8.6</td>
<td>21.2</td>
<td>9.0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>131</td>
<td>9.0</td>
<td>12.7</td>
<td>2.5</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>11</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>131</td>
<td>8.1</td>
<td>14.9</td>
<td>4.0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

Notes: n is the number of current year non survivors; mean, sd, sk, and px are the mean, standard deviation, skewness coefficient and x-th percentile of the non survivors previous year size distribution, respectively.
the skewness coefficient is due to the reduction in the number of surviving firms.\textsuperscript{2} However, when we look at the entering year distribution among survivors, in table 2, we conclude that it does not change significantly, meaning that, at least initially, survivors have accounted for much of the growth. On the other hand, from table 3, it seems that non survivors are mostly composed of those firms that did not benefit from growth.\textsuperscript{3} As a summary of the information in the tables we display, in figures 1 and 2, the kernel density estimates of the 1999 and 1988 size distributions among 1999 survivors.

Another important fact emerging from the age dependent firm dynamics is the different patterns of reallocation among young and old firms. We have calculated measures of gross job flows by age of the establishment for the manufacturing sector from Quadros de Pessoal, and show them in table 4.\textsuperscript{4} It is evident that young firms have high rates of job creation and job destruction, account for a significant part of all job creation and all job destruction, and have much higher net growth rates (both among all units and among continuing firms). This last fact makes a difference in the patterns of reallocation displayed by young versus old firms over the business cycle, an argument similar to that of Christopher Foote (1998), but now in

\begin{itemize}
  \item[\textsuperscript{2}]In these tables we did not control for false temporary exits, but will do it in the future. Because of this, the number of exiting firms implicit in table 1 differs from the (true) number of exiting firms in tables 2 and 3. This is especially relevant since there is a high incidence of false temporary exits among small firms, which might be accounting for some of the initial dislocation of the distribution to the right, when the distribution among exits does not seem to be much of a cause for it.
  \item[\textsuperscript{3}]In order to properly state this, we will need to get the initial year distribution of non survivors.
  \item[\textsuperscript{4}]Quadros de Pessoal contains very limited information on the age of establishments/firms. Basically, we defined the age to be that of the oldest worker of the establishment/firm, across all years, and have taken measures to avoid errors. In table 4, gross job flows are calculated for all establishments belonging to one-establishment firms and to those multieestablishment firms without any workers moving across establishments. However, the results are very similar to those obtained by considering only one establishment firms. In Eugénio Pinto (2005), we give a more complete description of the methodology and alternatives.
\end{itemize}
Figure 2: 1988 Firm Cohort: Kernel Density Estimate in 1988

Table 4: Gross Job Flows by Age in Manufacturing (1987-1999)

<table>
<thead>
<tr>
<th>Age</th>
<th>$esh$</th>
<th>$jc$</th>
<th>$jcc$</th>
<th>$jsh$</th>
<th>$jd$</th>
<th>$jdc$</th>
<th>$jds$</th>
<th>$sdr$</th>
<th>$c_{neti,net}$</th>
<th>$c_{rea,neti}$</th>
<th>$c_{erea,neti}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>7.5</td>
<td>44.0</td>
<td>14.3</td>
<td>29.4</td>
<td>14.4</td>
<td>5.0</td>
<td>10.1</td>
<td>0.24</td>
<td>0.91</td>
<td>0.90</td>
<td>-0.52</td>
</tr>
<tr>
<td>5-9</td>
<td>9.5</td>
<td>9.1</td>
<td>9.1</td>
<td>7.8</td>
<td>13.0</td>
<td>5.8</td>
<td>11.5</td>
<td>1.13</td>
<td>0.88</td>
<td>-0.13</td>
<td>0.78</td>
</tr>
<tr>
<td>10-</td>
<td>44.5</td>
<td>4.6</td>
<td>4.6</td>
<td>18.4</td>
<td>9.3</td>
<td>5.7</td>
<td>39.0</td>
<td>1.56</td>
<td>0.94</td>
<td>-0.55</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Notes: $esh$, $jsh$, and $jds$ are the shares of each age group in manufacturing employment, job creation, and job destruction, respectively; $jc$, $jcc$, $jd$, and $jdc$ are the rates of job creation, for all and for continuing establishments, job destruction, for all and for continuing establishments, respectively; $c_{x,y}$ is the correlation coefficient between $x$ and $y$, where $neti = jc - jd$, $rea = jc + jd$, and $erea = rea - |neti|$ are the rates of net job growth (for each age group and for all firms in manufacturing), of reallocation, and excess reallocation respectively.
a different context. We can see that job creation is much more volatile than job destruction for young firms, but the opposite occurs for old firms (and a similar situation occurs even if we consider only the flows among continuing units). This then means that reallocation is procyclical and excess reallocation is countercyclical among young firms, with the opposite occurring among older firms. This suggests that young and old firms are “forced” to adjust in different moments of the cycle. New firms, which have a tendency to grow, will be more active in auspicious macroeconomic periods, whereas old firms, with higher propensity to stagnate or decay, will be adjusting more in unfavorable macroeconomic moments.

We now make a summary of some reference empirical works. For the U.S. manufacturing sector, Dunne et al. (1989b) find that small and young (surviving) firms tend to grow faster and have higher failure rates than large firms do. Also, Dunne et al. (1989a) find that job destruction due to the scaling-down of firms activity increases with plant age, whereas job destruction due to plant exit and job creation due to the scaling-up of firms size decrease with plant age. For the Portuguese manufacturing sector, Mata and Portugal (1994) find that entering plants are small in absolute size and significantly smaller than existing plants, but the survivors exhibit significative positive growth rates (they are 50%/100% larger after 4/7 years). Mata et al. (1995) also conclude that the survival probability increases with the entry size (and current size) of firms, with the age of the firm, and the growth rate of the sector, and decreases with the entry rate to the sector.

In order to ascertain how generalized across countries are these facts, Bartelsman et. al. (2003) use a newly created dataset on firm demographics for 10 OECD countries, specifically built to make the cross-country analysis reliable. Their results confirm some of the stylized facts identified in previous studies. However, one important finding is that size distributions differ across countries and across industries, suggesting that technological differences are determinant, but also that country specific factors matter. In particular, the US displays a size distribution with much more dispersion, a smaller relative entry size (both due to smaller absolute entry sizes and to higher size of incumbent firms), and hazard rates that decline less steeply. The authors argue that this evidence could result from the structure of entry, exit and adjustment costs: entry costs can create disincentives for small firms to enter and then expand, and post-entry adjustment can be limited by the adjustment costs.

### 2.2 Some Theories

In this subsection we go over some of the main theoretical models that have been put forward to explain the above facts. Among them, the most well known is the passive learning model of Jovanovic (1982). In his model \textit{ex ante} identical firms learn about their \textit{ex post} average efficiency by observing, period by period, their performance. In each period, before they make their output decision, firms must decide if they remain in the industry, taking into account the opportunity cost of doing so. Because the reliability they put on the Bayesian posterior estimate of their productivity parameter is increasing with age, the exit frontier will be increasing over time. Therefore, inefficient firms gradually exit the market, as they learn that they are so, and efficient firms remain in the industry and adjust their employment level to be in accordance with their average productivity. The model is able to generate an increasing average size, and a decreasing failure rate with the age of a given cohort of
surviving firms. However, this model does not explain the dispersion in the entry size of firms, the dispersion in the exit size of firms, and the existence of a significant amount of small old firms.

Hugo Hopenhayn (1992) builds a model that describes an industry long run equilibrium with exit and entry of firms, and therefore it better describes the facts than explains them. In his model, firms learn their entry efficiency after paying the entry cost, so that there will be heterogeneity in the initial entry size of firms. Even though firms learn their productivity at entry, it evolves stochastically over time according to a Markov process, with firms deciding on exit before observing next period’s productivity. He then discusses conditions under which some of the above stylized facts will hold in the model, and produces some comparative statics analysis.

Richard Ericson and Ariel Pakes (1995) construct a model where the profitability of a firm is determined by the stochastic outcomes of the investment projects realized by the firm, by other firms in the industry and by the competitive environment in which the firm lives. Therefore, contrasting with Jovanovic (1982), this is an active learning model, in the sense that firms must spend resources in order to improve their relative position in the industry. They study the implications of the model for exit behavior, the size distribution of firms, and the evolution of the industry.

As an alternative explanation to the stylized facts on firms dynamics, Thomas Cooley and Vincenzo Quadrini (2001) consider a model with financial frictions and persistence in productivity shocks. In this model, the existence of a transaction cost on equity and a default cost on debt implies that equity and debt will not be perfect substitutes, and therefore size will depend positively on the amount of equity. Assuming that firms observe their efficiency before entering the industry, new firms will tend to be of high productivity and will borrow more intensively. This implies that they will face higher volatility in their performance. With this feature, the model is able to generate both the negative dependence on age (due to financial frictions) and the negative dependence on size (due to persistence of shocks) of growth, volatility of growth, job creation, job destruction and exit.

Cabral and Mata (2003) use the Quadros de Pessoal database to study the evolution of the firm size distribution among Portuguese manufacturing firms. They follow the 1984 cohort of entering firms until 1991, and conclude that at entry the distribution is very skewed to the right, but it gradually evolves towards a more symmetric distribution, approximately lognormal, essentially due to the evolution of the survivors in a given cohort. They argue that this evolution is more favorable to a theory of young financially constrained firms rather than a market selection mechanism.

In the next section, we explore the effect of nonconvex entry, regular and exit adjustment costs in a learning-about-efficiency model of industry dynamics. Eventually, those costs might be a source to explain some of the cross-country and cross-industry differences in the dynamic behavior of firms, and it might help to account for the differences in gross job flows between young and old firms. Our motivation is that adjustment costs in a learning environment provide incentives for young firms to be prudent. That is, young firms will take whatever decision that minimizes the probability that they will need to adjust in the future. For example, if firing costs are higher than hiring costs, young firms will want to start smaller, and as uncertainty decreases, they will adjust to their optimal long run dimension.
To our knowledge, this is the first time adjustment costs are being used in this context. After starting this project, we found the paper by Luís Cabral (1995) which is nearest to our idea. In his model, firms pay a proportional sunk capacity cost, that must be incurred any time the firm decides to increase its production capacity. Even though his model has a very simple structure, he argues that in a model of learning, where the conditional distribution of future profitability is monotonically increasing in today’s profitability (e.g. Jovanovic 1982), a proportional capacity cost would make small entering firms grow more fast than large entering firms. The reason is that because small entrants are those whose profitability signals were not so good, their exit probabilities are higher, and therefore they choose to invest more gradually. Consequently, they will exhibit higher growth rates than large entrants. However, the assumption that there is no fixed capacity cost, and that the capacity cost is completely sunk, does not seem the be very realistic.

In our model, we emphasize adjustment costs, instead of capacity costs, and are more interested in the age-growth relationship. The dynamics in our model is much richer, since we consider various kinds of adjustment costs, and discuss the influence of the distribution of efficiency and of the production function. We should also note that large entrants do register average insignificant increase in size, as can be seen in table 1, something that our model is able to account for.

Previous studies on the impact of nonconvex adjustment costs did not give it special importance in affecting such aspects as the average labor demand and the firm size distribution. For example, Samuel Bentolila and Giuseppe Bertola (1990) study the effects of linear firing and hiring costs for average labor demand and the dynamics of labor adjustment. They conclude that high firing costs do not have a significant effect on average labor demand, but make adjustments more sluggish. This would explain some facts in the 1970’s and 1980’s for the major European economies: with the oil price shocks and the slowly growing demand, firms did not fire so many workers as they wished to, but they were also reluctant to hire new workers even after a mild recovery started. Also, Hugo Hopenhayn and Richard Rogerson (1993), in an extension of Hopenhayn’s (1992) model to a general equilibrium setting, analyze the effects of a tax on job destruction on employment, productivity and welfare. They conclude that it reduces significantly employment, productivity and welfare. However, its effect on the size distribution of firms is not significant.

3 A Model of Learning with Nonconvex Costs

The model that follows borrows the Bayesian learning mechanism from Jovanovic (1982). We assume an industry with competitive output and input markets. In the absence of adjustment costs, the optimal decision is for the firm to choose its labor input, \( L \), in order to maximize current profits. Current profits are defined by

\[
\Pi (L, \theta) = F (L) \theta - wL,
\]

where \( F (\cdot) \) is the production function, exhibiting decreasing returns to labor, \( L \), \( \theta \) is a productivity shock, and \( w \) is the wage rate, which we will assume to be a known constant. The output price is the numeraire, so that \( w \) is the real wage rate. For now, we will not be interested in deriving the equilibrium of the model, so that we treat \( w \) as a given constant.
The stochastic process of $\theta$ is defined by\(^5\)

$$
\theta_{t+\tau} = \xi (\eta_{t+\tau}), \xi > 0, \xi > 0, \text{continuous, } \xi (-\infty) = \nu_1 \geq 0, \xi (\infty) = \nu_2 \leq \infty,
$$

and the firm knows that in every period $\eta$ is normally distributed with an unknown mean and known variance, that is,

$$
\eta_{t+\tau} = \mu + \varepsilon_{t+\tau}, \varepsilon_{t+\tau} \sim N(0, \sigma^2).
$$

With respect to $\mu$, the firm knows that it is a random draw from

$$
\mu \sim N(\bar{\mu}, \sigma^2_\mu).
$$

Because, in every period, the firm will observe $\theta_{t+\tau}$, but only after taking its exit and employment decisions, then it will use the information on $\{\eta_{t+\tau}\}$ in order to learn about its productivity parameter, $\mu$. Following Arnold Zellner (1971), a firm with age $\tau \geq 0$ has the following Bayesian posterior distribution for $\mu$\(^6\)

$$
\mu_{\tau, \text{po}} \sim N(Y_\tau, S_\tau),
$$

where

\begin{align}
Y_\tau & = \frac{\sum_{s=0}^{\tau-1} \eta_{t+s} \sigma^{-2} + \bar{\mu} \sigma^{-2}_\mu}{\sum_{s=0}^{\tau-1} \sigma^{-2} + \sigma^{-2}_\mu} = \frac{\bar{\eta}_{t,\tau} + \frac{1}{\tau} \bar{\mu} \sigma^{-2}_\mu}{1 + \frac{1}{\tau} \sigma^{-2}_\mu}, \bar{\eta}_{t,\tau} = \frac{1}{\tau} \sum_{s=0}^{\tau-1} \eta_{t+s}, \tag{1a}
\end{align}

\begin{align}
S_\tau & = \left(\sum_{s=0}^{\tau-1} \sigma^{-2} + \sigma^{-2}_\mu\right)^{-1} = \frac{1}{\tau \sigma^{-2} + \sigma^{-2}_\mu} = \frac{\sigma^2_\mu}{\tau \sigma^2_\mu + \sigma^2}. \tag{1b}
\end{align}

Before choosing how much labor to hire, the firm uses its updated information on $\mu$ to decide whether it will remain in the industry for one more period, in which case it expects to get a pay-off of $V$, or it will exit the industry and get the opportunity cost $W$. We now clarify the timing structure of the model. At the end of period $t + \tau - 1$, the firm observes its $\theta_{t+\tau-1} = \xi (\eta_{t+\tau-1})$, which it then uses, at the beginning of period $t + \tau$, to update its information set and calculate the (posterior) expectation of $\theta_{t+\tau}$, $\theta_{t+\tau}^\ast$. Given $\theta_{t+\tau}^\ast$ and $L_{t+\tau-1}$ the firm decides if it will continue in the industry, and, conditionally on staying, it makes its employment decision, $L_{t+\tau}$. The firm then produces and, by the end of period $t + \tau$, observes $\theta_{t+\tau}$. This process repeats itself until the firm decides to leave the industry.\(^7\)

\(^5\)From now on, we will assume $\tau$ to be the age of the firm.

\(^6\)We should mention that by using this setup we are not considering the existence of heterogeneity in entry size. One way to achieve this would be to assume that firms are randomly assigned an optimal capacity size, and that they receive a signal on their efficiency, which is somehow correlated with their true efficiency. However, they only learn about their true efficiency by using a Bayesian updating rule. This change would not hurt the main findings in the paper, and would be necessary in order to proceed with some calibration of the model. This is left for future work.

\(^7\)In this simple model we are not considering the possibility that as firms get older they might decay or become obsolete. This could be achieved by assuming some exogenous probability of becoming obsolete or decaying in efficiency. This is an extension that should be considered in any calibration experiment.
At any time \( t + \tau, \tau = 0, 1, \ldots \), conditionally on staying/entering the industry, the above optimization problem can be represented as

\[
V(\theta^*_t+\tau, \tau) = \max_{L_{t+\tau}} \{ \Pi(L_{t+\tau}, \theta^*_t+\tau) + \beta E_{t+\tau} \left[ \max \{ W, V(\theta^*_t+\tau+1, \tau + 1) \} \right] \}
\]

(2)

where \( E_{t+\tau}(\cdot) \equiv E(\cdot | \Omega_{t+\tau}) \), \( \Omega_{t+\tau} \equiv \{ \eta_{t+s} \}_{s=0}^{\tau-1} \), represents the conditional expectation operator given the information at \( t + \tau \), \( \Omega_{t+\tau} \), which can be summarized by \((\theta^*_t+\tau, \tau) \), where \( \theta^*_t+\tau = E_{t+\tau}(\theta_{t+\tau}) \). In this optimization there are no intertemporal linkages, namely, because we are omitting adjustment costs, so that, if \( \theta^*_t+\tau \) is above some (time-dependent) exit threshold \( \theta^0_{t+\tau} \), optimal employment will be chosen to maximize current period profits. For values of \( \theta^*_t+\tau \) below this threshold, the firm finds it optimal to leave the industry, because its expected profitability will not make up for the opportunity cost. The increasing confidence the firm puts in \( \theta^*_t+\tau \), as it grows older, implies that the exit threshold will be increasing with age. This is the main element of the model, and lies behind the age effects argument used by Jovanovic (1982): the tendency of the size distribution and the survival probability to increase with age.

For some of the results below we will assume that \( F \) is Cobb-Douglas with decreasing returns to scale, i.e., \( F(L) = L^\alpha, \alpha \in (0, 1) \). In this case, optimal decisions are

\[
L^*_t+\tau = L_{t+\tau}(\theta^*_t+\tau, \tau) = \begin{cases} \bar{L}(\theta^*_t+\tau) \equiv \left( \frac{w \alpha \theta^*_t+\tau}{\omega} \right)^{\frac{1}{1-\alpha}}, & \text{if } \theta^*_t+\tau \geq \theta^0_{t+\tau}, \\ 0, & \text{if } \theta^*_t+\tau < \theta^0_{t+\tau}. \end{cases}
\]

(3)

Note that under the Cobb-Douglas specification, with \( \alpha \in (0, 1) \), optimal employment is a convex function of \( \theta^*_t+\tau \). This implies that surviving firms will grow over time, simply because of Jensen’s inequality and because \( \theta^*_t+\tau \) is a Martingale: \( E_{t+\tau} \left[ \bar{L}(\theta^*_t+\tau+1) \right] > \bar{L} \left[ E_{t+\tau}(\theta^*_t+\tau+1) \right] = \bar{L}(\theta^*_t+\tau) \). However, this is not a result that holds in general.9

We now introduce nonconvex adjustment costs into the model, allowing entry and exit costs to differ from regular adjustment costs. An essential part of the model is the fixed entry cost, because it diminishes the intensity of the selection process, avoiding the extreme situation where entry would be so high that only the highest productivity firms would remain in the industry. The regular adjustment cost function is then

\[
C^S(L_{t+\tau}, L_{t+\tau-1}) = \begin{cases} H_f + H_p(L_{t+\tau} - L_{t+\tau-1}), & \text{if } L_{t+\tau} > L_{t+\tau+1}, \\ 0, & \text{if } L_{t+\tau} = L_{t+\tau+1}, \\ F_j + F_p(L_{t+\tau-1} - L_{t+\tau}), & \text{if } L_{t+\tau} < L_{t+\tau+1}, \end{cases}
\]

\[8\text{Throughout the text we will keep the notation } \bar{L}(x) \text{ to denote } (\frac{\alpha x}{\omega})^{\frac{1}{1-\alpha}}.\]

\[9\text{In general, we have}
\]

\[
\bar{L}''(\theta^*) = \left( \frac{F''(\bar{L}) F' (\bar{L}) - 2 F''(\bar{L})^2}{F''(\bar{L})^2} \right) L'(\theta^*), \quad F''(\bar{L}) < 0, \quad L'(\theta^*) > 0,
\]

whose sign depends on \( F''(\bar{L}) \). Therefore, if decreasing returns to labor do not decrease too fast, that is, \( F''(\bar{L}) < 2 F''(\bar{L})^2 / F'(\bar{L}) \), then we will have \( \bar{L}''(\theta^*) < 0 \). For example, when \( F(L) = \ln(L) \), then \( \bar{L}''(\theta^*) = 0 \).
and the entry and exit adjustment cost functions are

\[
C^I (L_t) = I_f + I_p L_t, \quad I_f > 0, \\
C^O (L_{t+\tau}) = O_f + O_p L_{t+\tau}.
\]

With adjustment costs, the optimization problem now becomes,

\[
V^S (L_{t+\tau-1}, \theta^*_{t+\tau}, \tau) = \max_{L_{t+\tau}} \left\{ \left[ \Pi (L_{t+\tau}, \theta^*_{t+\tau}) - C^S (L_{t+\tau}, L_{t+\tau-1}) \right] \\
+ \beta E_{t+\tau} \left[ \max \left\{ V^O (L_{t+\tau}), V^S (L_{t+\tau}, \theta^*_{t+\tau+1}, \tau + 1) \right\} \right] \right\}, \tag{4}
\]

for all periods after entry (\(\tau \geq 1\)) in which the firm remains in the industry, and,

\[
V^I = \max_{L_t} \left\{ \left[ \Pi (L_t, \theta^*_t) - C^I (L_t, L_{t+\tau-1}) \right] + \beta E_t \left[ \max \left\{ V^O (L_{t+\tau}), V^S (L_t, \theta^*_{t+1}, 1) \right\} \right] \right\}, 
\tag{5}
\]

for the entry period, where \(V^O\), the value of exiting, is defined as

\[
V^O (L_{t+\tau}) = W - C^O (L_{t+\tau}).
\]

The superscript indexes \(S\), \(I\), and \(O\) stand for “staying”, “entering”, and “exiting” the industry. Note that contrary to the model without adjustment costs, the previous period employment is a state variable for the current period optimization problem. Also, in \(V^I\) and in \(V^O\) the cost of entry and the cost of exit, respectively, are taken into account.

This problem is very hard to solve analytically, since the objective function is not continuous (implying nonconcavity) and is not differentiable at the frontiers between adjustment and non adjustment Therefore, we discuss briefly how the problem can be mathematically solved. First note that

\[
V^S (\cdot) = \max \left\{ V^{SD} (\cdot), V^{SN} (\cdot), V^{SU} (\cdot) \right\},
\]

where the superscript indexes \(SD\), \(SN\), and \(SU\) stand for “staying and adjusting downwards”, “staying and not adjusting”, and “staying and adjusting upwards”. That is, the firm chooses between those three discrete actions, after considering the optimal employment decision in each of the them. Namely,

\[
V^{SD} (L_{t+\tau-1}, \theta^*_{t+\tau}, \tau) = \max_{L_{t+\tau} \leq L_{t+\tau-1}} \left\{ \left[ \Pi (L_{t+\tau}, \theta^*_{t+\tau}) - F_f - F_p (L_{t+\tau-1} - L_{t+\tau}) \right] \\
+ \beta E_{t+\tau} \left[ \max \left\{ V^O (L_{t+\tau}), V^S (L_{t+\tau}, \theta^*_{t+\tau+1}, \tau + 1) \right\} \right] \right\}, 
\]

\[
V^{SN} (L_{t+\tau-1}, \theta^*_{t+\tau}, \tau) = \Pi (L_{t+\tau-1}, \theta^*_{t+\tau}) \\
+ \beta E_{t+\tau} \left[ \max \left\{ V^O (L_{t+\tau-1}), V^S (L_{t+\tau-1}, \theta^*_{t+\tau+1}, \tau + 1) \right\} \right], 
\]

\[
V^{SU} (L_{t+\tau-1}, \theta^*_{t+\tau}, \tau) = \max_{L_{t+\tau} \geq L_{t+\tau-1}} \left\{ \left[ \Pi (L_{t+\tau}, \theta^*_{t+\tau}) - H_f - H_p (L_{t+\tau} - L_{t+\tau-1}) \right] \\
+ \beta E_{t+\tau} \left[ \max \left\{ V^O (L_{t+\tau}), V^S (L_{t+\tau}, \theta^*_{t+\tau+1}, \tau + 1) \right\} \right] \right\}.
\]
Before presenting the optimality conditions associated with $V^S$ and $V^I$, for given $(L_{t+\tau-1}, \tau)$ we partition the state-space associated with $\theta^*_t$, $\Theta$, into regions of exit, $\Theta^O$, downward adjustment, $\Theta^{SD}$, non-adjustment, $\Theta^{SN}$, and upward adjustment, $\Theta^{SU}$:

$$\Theta^O (L_{t+\tau-1}, \tau) = \{ \theta : V^O (L_{t+\tau-1}) > V^S (L_{t+\tau-1}, \theta, \tau) \},$$

$$\Theta^{SD} (L_{t+\tau-1}, \tau) = \{ \theta : V^{SD} (L_{t+\tau-1}, \theta, \tau) > V^{SN} (L_{t+\tau-1}, \theta, \tau), \}
V^{SD} (L_{t+\tau-1}, \theta, \tau) > V^O (L_{t+\tau-1}) \},$$

$$\Theta^{SN} (L_{t+\tau-1}, \tau) = \{ \theta : V^{SN} (L_{t+\tau-1}, \theta, \tau) \geq V^{SD} (L_{t+\tau-1}, \theta, \tau),
V^{SN} (L_{t+\tau-1}, \theta, \tau) > V^{SU} (L_{t+\tau-1}, \theta, \tau),
V^{SN} (L_{t+\tau-1}, \theta, \tau) > V^O (L_{t+\tau-1}) \},$$

$$\Theta^{SU} (L_{t+\tau-1}, \tau) = \{ \theta : V^{SU} (L_{t+\tau-1}, \theta, \tau) > V^{SN} (L_{t+\tau-1}, \theta, \tau),
V^{SU} (L_{t+\tau-1}, \theta, \tau) > V^O (L_{t+\tau-1}) \}.$$

Given this, we prove in appendix A the following proposition\(^{10}\)

**Proposition 1** In problem (4), the firm adjusts upwards if $V^{SU} > V^{SN}$, in which case optimal employment is determined by

$$[\alpha L_{t+\tau-1}^* \theta_{t+\tau}^* - w] + \sum_{s=1}^{\infty} E_{t+\tau} \beta^s \left\{ \chi^*_{t+\tau+s} (-O_p) + \hat{\chi}^*_{t+\tau+s} \left[ \alpha L_{t+\tau+s}^* \theta_{t+\tau+s}^* - w \right] \right\} = H_p, \hspace{1cm} (6)$$

and it adjusts downwards if $V^{SD} > V^{SN}$, in which case optimal employment is determined by

$$[\alpha L_{t+\tau-1}^* \theta_{t+\tau}^* - w] + \sum_{s=1}^{\infty} E_{t+\tau} \beta^s \left\{ \chi^*_{t+\tau+s} (-O_p) + \hat{\chi}^*_{t+\tau+s} \left[ \alpha L_{t+\tau+s}^* \theta_{t+\tau+s}^* - w \right] \right\} = -F_p, \hspace{1cm} (7)$$

where $L_{t+\tau+s}^*$ and $\chi^*_{t+\tau+s}$ are the optimal policy functions for the exit and the employment decisions in period $t + \tau + s$. $\chi^*_{t+\tau+1}$, $\hat{\chi}^*_{t+\tau+1}$, and $\tilde{\chi}^*_{t+\tau+1}$ are defined as

$$\chi^*_{t+\tau+1} (L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 1) = \begin{cases} 1, & \text{if } \theta_{t+\tau+1}^* \in \Theta^O (L_{t+\tau}, \tau + 1), \\ 0, & \text{otherwise}. \end{cases}$$

and

$$\tilde{\chi}^*_{t+\tau+s} = \chi^*_{t+\tau+s} \prod_{j=1}^{s-1} (1 - \chi^*_{t+\tau+j}) = \begin{cases} 1, & \text{if } \chi^*_{t+\tau+j} = 0, j \in \{1, \ldots, s-1\}, \chi^*_{t+\tau+s} = 1, \\ 0, & \text{otherwise}, \end{cases}$$

$$\hat{\chi}^*_{t+\tau+s} = \prod_{j=1}^{s} (1 - \chi^*_{t+\tau+j}) = \begin{cases} 1, & \text{if } \chi^*_{t+\tau+j} = 0, \text{for } j \in \{1, \ldots, s\}, \\ 0, & \text{otherwise}. \end{cases}$$

\(^{10}\)It is obvious that in the following formulas we could have written, more generally, the marginal revenue product as $F' (L_{t+\tau}) \theta_{t+\tau}^*$, instead of $\alpha L_{t+\tau}^* \theta_{t+\tau}^*$. 
The first two conditions are marginal conditions, similar to the smooth pasting conditions in the \((S, s)\) model literature, and they state that if the firm adjusts then the marginal adjustment cost must equal the expected present discounted value of the marginal revenue product for all future periods that the firm is still in the industry, minus the increase in the exit cost when the firm decides to exit. This is the discrete-time analog of the continuous-time result presented in Steven Nickell (1986), Bentolila and Bertola (1990), and Giuseppe Bertola (1992), adequately adjusted for the fact that now we also have an exit decision in each period. The firm will not adjust if the marginal cost of adjustment exceeds its marginal benefit for the first unit of adjustment. The two last conditions are value function conditions, similar to the value matching conditions in the \((S, s)\) model literature, and they state that adjustment will only occur when it has a better pay-off than non adjustment. Finally note that \(\chi^*_t+\tau+s\) equals one if the firm decides to stay in the industry in period \(t+\tau+s\), \(\tilde{\chi}^*_t+\tau+s\) assumes the value of one when the firm has remained in the industry until period \(t+\tau+s-1\), but decides to exit in period \(t+\tau+s\), and \(\hat{\chi}^*_t+\tau+s\) equals one when the firm does not exit also in \(t+\tau+s\).

Note that when there are fixed adjustment costs, contrary to what occurs in Nickell, Bentolila and Bertola, and Bertola, the fact that at the current employment level, \(L^*_t+\tau-1\), the marginal benefit of adjustment exceeds the marginal cost of adjustment does not imply that it is optimal for the firm to adjust, because the total benefit from adjustment must also exceed the total cost of adjustment (where the fixed cost has an impact). It is from this that we will be able to ascertain the impact of fixed adjustment costs on the optimal employment policy. We now present the optimality conditions associated with the entry period (see 10).

**Proposition 2** In problem (5) the firm enters the industry if \(V_I \geq W\), in which case optimal employment at entry is determined by

\[
[\alpha L^*_t-\theta^*_t-w] + \sum_{s=1}^{\infty} E_t \beta^s \left\{ \tilde{\chi}^*_t+s (-O_p) + \tilde{\chi}^*_t+s \left[ \alpha L^*_t-\theta^*_t+s-w \right] \right\} = I_p.
\]

**Proof.** Obvious from proposition 1. ■

The first condition says that the firm will enter only if it expects to be at least compensated for the opportunity cost. However, in general equilibrium we should have an equality, because while \(V_I > W\) new firms would enter the industry and would raise the real wage until an equality would hold. The following corollary of proposition 1 will also be useful for us to make qualitative statements about the effects of adjustment costs (see 10).

**Corollary 1** In problem (4), the marginal benefit of one additional unit of labor, that is, the LHS of expressions (6) and (7), can be recursively represented as

\[
MB \left( L^*_t, \theta^*_t, \tau \right) = \left( \alpha L^*_t-\theta^*_t-w \right) + \beta E_{t+\tau} \left[ \chi^*_t+\tau+1 (-O_p) \right. \left. + \left(1-\chi^*_t+\tau+1 \right) MB \left( L^*_t, \theta^*_t+1, \tau+1 \right) \right].
\]
Proof. See appendix A. ■

Even though the results in propositions 1 and 2, and corollary 1 do not enable us to solve the model analytically, they will be useful for us to make qualitative statements about the effects of adjustment costs. As we have seen above, when there are no adjustment costs, the optimal employment policy will always be given by 3, so that growth is a by-product of a selection mechanism: those firms that are inefficient, and therefore small, exit, while those firms that are efficient grow in size.\footnote{As we noted before, another growth effect, which does not hold in general, is due to the convexity of the employment decision rule in $\theta^*$.} We will now be considering different possibilities and combinations of adjustment costs, and see that these adjustment costs generate alternative sources of growth.

First consider exit adjustment costs in isolation. The fixed exit cost by itself will not affect the employment decision. It will just make the exit alternative less attractive, implying a decrease in the exit threshold, and, as a general equilibrium effect, a decrease in the real wage. Therefore, in the new equilibrium, we should expect an increase in initial size, but a slower growth rate among survivors due to less intense selection. This can enable us to see the contribution of adjustment costs for firm growth, independently of the selection mechanism.

The proportional exit cost will, on the contrary, affect the incentives to hire workers. The fact that larger firms will have higher costs upon exit creates an incentive to start smaller than otherwise, and gradually adjust to the decision that would have been taken if $O_p = 0$, as the firm starts realizing that the prospects of having to leave the industry are very tight. Note that in the case where there is just an exit proportional adjustment cost, in all periods, expression (8) will have to equal zero. This then implies that the optimal employment policy will be given by

$$ L^{*\tau}_{t+\tau} = \frac{\alpha \theta^*_{t+\tau}}{w + O_p \beta \Pr (\chi^*_{t+\tau+1} = 1)} \left( \frac{1}{1 - \alpha} \right). $$

Now take into account that, among surviving firms, $\Pr (\chi^*_{t+\tau+1} = 1)$ will be decreasing over time, because the reduction in uncertainty about $\theta^*_{t+\tau+1}$, given $\theta^*_{t+\tau}$, dominates the increase in the exit threshold. But, besides this, as time goes by, we are being left with only those firms that are really efficient and will not exit the market. Both of these effects will imply an incentive for firms to start smaller and gradually increase their size over time, being therefore an alternative explanation for growth.\footnote{Similarly to \ref{footnote:convexity}, we could have used exactly the same arguments without making use of the Cobb-Douglas production function. We are, and will be, using it to make the argument more clear.}

Consider now entry adjustment costs in isolation. The fixed entry adjustment cost mainly affects the real wage, and therefore the exit threshold and how much growth there will be due to selection of more efficient firms. In fact, from proposition 2, we know that $V_I \geq W$, and that the higher is $I_f$, the smaller is $V_I$. Therefore, in response to an increase in $I_f$, there will be less entry to the industry, a decrease in the real wage, and a decrease in the exit threshold. Now consider the case of a proportional entry cost. In the absence of regular adjustment costs, this will only imply a reduction in the entry size of firms, and an immediate upward adjustment to employment path that would have occurred in the absence of any adjustment.
costs. We can see this by using corollary 1 in proposition 2 to get
$$MB(L^*_t, \theta^*_t, 0) = (\alpha L^*_t^{\alpha - 1} \theta^*_t - w) + \beta \Pr(\chi^*_{t+1} = 0) E_t [MB(L^*_{t+1}, \theta^*_{t+1}, 1) \mid \chi^*_{t+1} = 0] = I_p,$$
where we are considering $O_p = 0$. But, because there are no regular adjustment costs, then $MB(L^*_{t+1}, \theta^*_{t+1}, s) = 0$, for all $s \geq 1$. Therefore, initial employment will be
$$L^*_t = \left(\frac{\alpha \theta^*_t}{w + I_p}\right)^{\frac{1}{1-\alpha}},$$
and, for all future periods, we will have $L^*_{t+s} = \bar{L}(\theta^*_{t+s})$. This effect has nothing new, since the same would occur in a model without learning. However, the interaction of the entry costs with the other costs might be interesting. We dive into this below.

Finally, we try to figure out the impact of regular adjustment costs. If we consider no proportional firing cost, the consequence of the fixed firing cost is that even thought the marginal benefit of firing a worker might be positive, it must be sufficiently positive to make the cost of adjustment smaller than the benefit. Therefore, in this case, (8) will often be negative, but never positive. This implies that the optimal size will be given by
$$L^*_t = \left(\frac{\alpha \theta^*_t}{w - \beta \Pr(\chi^*_{t+1} = 0) E_t [MB(L^*_{t+1}, \theta^*_{t+1}, 1) \mid \chi^*_{t+1} = 0]}\right)^{\frac{1}{1-\alpha}},$$
which is smaller than otherwise. From here, we see that the mass of $\theta^*_t$’s for which future $MB$’s will be zero (meaning hiring workers) will be much higher than otherwise, which will cause employment to grow faster. Obviously this process will eventually cease, as the firm realizes its efficiency level.

In the case of a proportional firing cost, but no fixed firing cost, we will have something similar to the above result. The difference will be that those firms who decide to reduce employment will do it less intensely, since the higher the reduction the higher the cost will be. To see this, note that at entry we will have the same expression as for the fixed firing cost case, but now we know that $MB(L^*_{t+1}, \theta^*_{t+1}, 1) \in (-F_p, 0]$. For all future periods, if the firm decides to adjust downwards, then it will do according to
$$L^*_{t+\tau} = \left(\frac{\alpha \theta^*_{t+\tau}}{(w - F_p) - \beta \Pr(\chi^*_{t+\tau+1} = 0) E_{t+\tau} [MB(L^*_{t+\tau+1}, \theta^*_{t+\tau+1}, \tau + 1) \mid \chi^*_{t+\tau+1} = 0]}\right)^{\frac{1}{1-\alpha}},$$
which is higher than otherwise because of $F_p$ in the denominator. Of course, when viewed in isolation, the analysis for the hiring adjustment costs will be analogous, but in the opposite direction. Firms will tend to start higher, and grow at a slower rate than otherwise, since the selection effect is partially counteracted by the fact that firms will want to start higher.

We now analyze what interesting combinations of adjustment costs might occur. We are mainly interested in the case of symmetry between entry/exit and regular costs. Consider first fixed entry and hiring costs, e.g., $I_{f,2} - I_{f,1} = H_f$ (i.e., we are comparing two situations where the fixed entry cost increases by the amount of the hiring fixed cost). At entry, we
will have \( V_t \geq W \), and \( MB(L_t^*, \theta_t^*, 0) = 0 \), so that initial employment will be given by (9). Because after entry we will tend to have \( MB > 0 \), then firms will be starting higher than otherwise, in order to avoid future hirings.\(^{13}\) Also, in the next few periods, hirings will be less frequent than firings, even though when they occur they will be higher than otherwise. Therefore, we expect that the final effect will be firms entering larger and growing less over time.

We examine now the case where proportional entry and hiring costs are symmetric, i.e., \( I_p = H_p \). At entry, we must have \( MB(L_t^*, \theta_t^*, 0) = H_p \), so that,

\[
L_t^* = \left( \frac{\alpha \theta_t^*}{(w + H_p) - \beta \Pr(\chi_{t+1}^* = 0) E_t[MB(L_{t+1}^*, \theta_{t+1}^*, 1) | \chi_{t+1}^* = 0]} \right)^{1/(1-\alpha)}.
\]

Because after entry we must have \( MB \in [0, H_p) \), then it is obvious that, in order to reduce the chances of future firings, firms will enter smaller. Now, if they hire in the future, they will do it similarly to (10), whereas, if they fire, they will do it similarly to (9). From this we derive three conclusions. First, firings will be less frequent than hirings. Second, when there is hiring, \( \Pr(\chi_{t+1}^* = 0) \) is close to 1, and \( E[MB(L_{t+1}^*, \theta_{t+1}^*, 1) | \chi_{t+1}^* = 0] \) is close to \( H_p \), so that firms will hire more intensely. Third, when there is firing, both these two elements will be close to 0, so that firms will fire more intensely. In the end, we expect that the first effect will be dominating, so that growth will be higher.

A similar analysis in the case of symmetric exit and firing costs will reveal that both for fixed and proportional adjustment costs, we will have firms starting smaller and growing faster. simply because firms want to avoid future firings.

In conclusion, the only scenario that will cause firms to start higher and grow slower will be the case of high fixed hiring (and entry) nonconvex adjustment costs. This is certainly a reasonable explanation for the fact found in Bartelsman et al. (2003), that in the US and Canada, firms tend to enter smaller and grow faster than in European countries. This could suggest that small entry hiring fixed costs would make selection much more intense. However, we believe this is not a necessary implication. The differences in the selection rate will of much small order than the differences in entry size and growth, suggesting that most of the differences will be explained by the effect of adjustment costs in firms intertemporal optimization problem.

Before finishing this section, we should also mention that even when adjustment costs are symmetric, and we do not consider the effect of entry or exit costs, and the effect of selection, the fact that the distribution of productivity shocks is skewed to the right and that \( \alpha > 1/2 \), will create incentives for firms to start smaller and grow faster than otherwise. The idea is that with uncertainty concerning its efficiency, firms will want to be situated around points where the probability that it will need to adjust in future periods is smaller. When the distribution is skewed to the right firms know that most probably their efficiency will be below average, and when \( \alpha > 1/2 \), firms know that overemployment brings a higher decrease in profits than underemployment. We know that, at least, the first assumption is widely observed/held in the literature (see, for example, Bee Aw, Xiaomin Chen and Mark Roberts 1997). Therefore, under each of those two circumstances, when uncertainty is high

\(^{13}\)Another effect, in the same direction, is that the increase in \( I_t \) will make \( w \) decrease.
firms will want to underhire, but as uncertainty is resolved they will gradually increase their labor force. We leave the explanation of these statements for an example in appendix B.

4 Model Simulation Under Finite Learning Horizon

The most important obstacle we face in simulating the infinite learning horizon model is that $V^S$ depends on the firm age. This prevents us to use an iterative method that in each iteration provides some converging approximation to the value function or the policy function. Therefore, we follow the suggestion of Lars Ljungqvist and Thomas Sargent (2000) and consider an approximation where at some age $\theta$ parameter is reasonable in empirical studies (e.g. Aw et al. 1997). Given this assumption, we have the following proposition concerning the transition law for the $\theta^*$’s.

**Proposition 3** Let $\theta_{t+\tau} = \exp \{ \eta_{t+\tau} \}$, $\eta_{t+\tau} = \mu + \varepsilon_{t+\tau}$, $\varepsilon_{t+\tau} \sim \text{iidN}(0, \sigma^2)$, $\mu \sim \text{N}(\bar{\mu}, \sigma^2)$, $\{\varepsilon_{t+s}\}_{s \geq 0}$ independent from $\mu$. Then, the posterior distribution of $\theta_{t+\tau+j}$ ($j \geq 0$) given information $\Omega_{t+\tau} = \{\eta_{t+s}\}_{s=0}^{\tau-1}$, if $\tau < T$, and $\Omega_{t+\tau} = \mu$ if $\tau \geq T$, is

$$
\theta_{t+\tau+j} | \Omega_{t+\tau} \sim \text{log N} \left( Y_{\tau}, S_{\tau} + \sigma^2 \right),
$$

where, for $\tau < T$, $Y_{\tau}$ and $S_{\tau}$ are defined in (1), and, for $\tau \geq T$, $Y_{\tau} = \mu$, $S_{\tau} = 0$. Let $\theta^*_{t+\tau} = E(\theta_{t+\tau} | \Omega_{t+\tau}) = E(\theta_{t+\tau} | \theta^*_{t+\tau}, \tau)$. Then the distribution of $\theta^*_{t+\tau+j}$ ($j \geq 1$) given $(\theta^*_{t+\tau}, \tau)$ is

$$
\theta^*_{t+\tau+j} | (\theta^*_{t+\tau}, \tau) \sim \text{log N} \left( \ln(\theta^*_{t+\tau}) - \frac{1}{2} (S_{\tau} - S_{\tau+j}), S_{\tau} - S_{\tau+j} \right).
$$

**Proof.** See appendix A. □

Given a choice for $T$, the next proposition analyzes the properties of the optimization problem after $\mu$ is revealed to the firm.

**Proposition 4** If $\mu$ is known to the firm at period $t + T$, then all adjustments are made at period $t + T$, and the firm will not change its exit and employment decisions after that period. This means that

$$
V^S (\theta^*_{t+T}, L_{t+T-1}, T) = \max_L \left\{ \frac{1}{1 - \beta} \Pi(L, \theta^*_{t+T}) - C^S (L, L_{t+T}) \right\},
$$

$$
L^*_{t+s} = L(\theta^*_{t+T}, L_{t+T-1}, T), s \geq T, \chi^*_{t+T} = 1 [V^S (\theta^*_{t+T}, L_{t+T-1}, T) < V^O (L_{t+T-1})].
$$

**Proof.** See appendix A. □

This result enable us to simplify significantly the computational algorithm, since we basically have a finite horizon dynamic programming problem. In appendix C we present some details concerning the computational algorithm we use. For now, we just mention how to iteratively compute the densities of firms across the state-space in each period.
Proposition 5 Consider a generic period $t + s$, and let $NE^{s-1}$ represent the event that a firm has survived through period $t + s - 1$, and it has not considered yet if it will remain in period $t + s$, and let $NE_{s-1}$ represent the event that a firm will produce in period $t + s - 1$. That is,

$$NE_{s-1} = \{ (\theta^*_t, L_t) : \theta^*_t \geq \theta^*_{t+1} (L_{t+1}) \},$$

$$NE^{s-1} = \{ (\theta^*_t, \ldots, \theta^*_t, L_{t+1}, \ldots, L_{t+s-2}) : \theta^*_t \geq \theta^*_{t+1} (L_t), \ldots, \theta^*_t \geq \theta^*_{t+s-1} (L_{t+s-2}) \}.$$

where $NE^{s-1} = NE^{s-2} \cup NE_{s-1}$. Then, for $s \geq 2$

$$f (\theta_{t+s}, L_{t+s-1} \mid NE^{s-1})$$

$$= \left[ \int_{NE_{s-1}} f (\theta_{t+s-1}, L_{t+s-2} \mid NE^{s-2}) d\theta_{t+s-1} dL_{t+s-2} \right]^{-1}$$

$$\times \int_{NE_{s-1}} f (\theta^*_{t+s} \mid \theta^*_{t+s-1}, s-1) 1 [L_{t+s-1} = L_{t+s-1} (\theta^*_{t+s-1}, L_{t+s-2})]$$

$$f (\theta^*_{t+s-1}, L_{t+s-2} \mid NE^{s-2}) d\theta^*_{t+s-1} dL_{t+s-2}$$

Proof. See appendix A. ■

We are now ready to present some simulation results. We have chosen the following base specification for the model

$$\alpha = 0.66, \beta = 0.95, W = 100, w = 10, \bar{\mu} = 2.5, \sigma^2 = 0.07, \sigma^2 = 0.1$$

$$I_f = 50, I_p = 0, F_f = 0, F_p = 0, H_f = 0, H_p = 0, O_f = 0, O_p = 0.$$ 

First of all, we should mention that under the specifications we have used, not considering adjustment costs, there are two sources of growth. The first is due to the convexity of $L(\cdot)$ and the martingale property of $\theta^*$, and the second is due to the selection mechanism. Because the first does not need to hold in general, in the exercises below we would like to know the contribution of the selection mechanism. Therefore, in the first simulation we assume $O_f = 80$. The results are presented in figure 1. We can see that such a high exit cost shuts down exit completely. We also see that the contribution of selection is higher the smaller the age of firms.

We start by analyzing the impact of firing costs. To have an idea of the potential interaction between the selection mechanism and adjustment costs, we compare the case with adjustment costs with two alternatives: first, the base case, and, second, a case with no adjustment costs and no selection. For case 1, we set $F_f = 5, F_p = 3, O_f = 5$, and $O_p = 3$ (see figure 2), and, for case 2, we set $F_f = 5, F_p = 3, O_f = 80$, and $O_p = 3$ (see figure 14).

The discrete jumps that occur at period 25 are due to the fact that then firms get to know their true productivity parameter, and therefore adjust accordingly. The jump would have been smaller if we had use a larger number of periods. Also, $p$ is considered to be the price. We have used $p$ instead of $W$ as the variable to be changed in order to obtain the general equilibrium condition $V' = W$. See appendix C for further details.
The results show that in the first situation, growth is increased, even though selection is smaller, and that firms will start smaller and gradually converge to the size consistent with the situation without firing costs. From figure 3, we can see that, if anything, the interaction with the selection mechanism goes in the way of diminishing the impact of firing costs.

For the case of hiring costs, if we considered $I_{f,2} - I_{f,1} < H_f$, $I_p < H_p$, firms would be interested in starting with a size larger than otherwise, because by doing so they would be saving in hiring costs in the second period. Therefore, we assume instead $I_f = 55$, $I_p = 3$, $H_f = 5$, and $H_p = 3$, with results in figure 4. As should be expected, the fixed entry cost makes firms to start higher than otherwise. Most of the differences in growth rates is due to this effect, even though the selection is much more harsh in case 1. The exit barriers are much smaller now mainly because of the increase in price, which also explains the smaller selection. If we had shut down exit, the effect of hiring costs would be slightly more significant. However, our results would have been completely changed, namely, firms would have started smaller and grown faster, if instead we had assumed $H_f = 0$, and $I_f = 50$. As we mentioned in the previous section, an increase in fixed entry costs is the only reason why growth should be smaller.

We now consider that firing costs are higher than hiring costs, entry costs equal hiring costs, and exit costs equal firing costs, namely, $F_f = O_f = 5$, $F_p = O_p = 3$, $I_f = 52$, $H_f = 2$, $I_p = H_p = 1$. From figure 5 we see that growth is higher than without adjustment costs, essentially due to firms starting smaller, even though the price level is higher and selection is diminished. The problem associated with a lower growth rate at entry, is due to the selection mechanism, and the very high fixed entry cost, when compared to the proportional entry cost (see figure 6). More interestingly, even in the opposite case, that is, when $F_f = O_f = 2$, $F_p = O_p = 1$, $I_f = 55$, $H_f = 5$, $I_p = H_p = 3$, we would also have higher growth with adjustment costs (see figure 7).

From these exercises, we can think of two situations, netting out the contributions associated with the convexity of $\bar{L}$, and the selection mechanism. In the first case, fixed entry and hiring costs would be the dominating force, so that firms would tend to enter larger and growth would be mainly due to a selection mechanism. In the second case, fixed hiring and entry costs would be dominated by the other adjustment costs. This would make firms start smaller, and grow faster, so that growth of survivors would be the main force behind the increase in size among the cohort of entrants.

5 Plans for Future Work

We believe our argument has multiple possibilities of being tested empirically. In work under way we are trying to use it to explain differences in entry size and growth of firms across countries and industries. A good source of data for this work is the recent OECD firm-level dataset (see Bartelsman et. al. 2003), and the Quadros de Pessoal database. Another possibility would be to use micro data in order to structurally estimate the adjustment parameters. That would force us to change some aspects of the model we have simulated. We would need to introduce, namely, an heterogeneous entry size, a decay and obsolescence exogenous probabilities, and some aggregate industry demand function. We do think that
both of these projects are interesting of their own.

6 Conclusion

In this paper, we incorporated nonconvex adjustment costs into a standard model of firm dynamics with Bayesian learning due to Jovanovic (1982). The conclusion was that except for the case where entry and hiring fixed costs are significant, growth is significantly generated by the increase in size of survivors, and not by the elimination of nonefficient small firms. This might open an empirical question of knowing whether, for the stochastically increasing firm size distribution on age, it is more determinant the growth of survivors or, instead, the exit of small firms.

Cabral and Mata (2003) found that, among the Portuguese manufacturing firms, the most important contribution for the evolution of a given cohort’s size distribution was the first effect. To test if firms start small because they are financially constrained, they used the entrepreneur’s age as a predictor for size, arguing that older entrepreneurs tend to be richer. They concluded that young entrepreneurs tend to have smaller firms, giving credit to their theory. However, they remark: “An alternative interpretation is that entrepreneur’s age is a good proxy for previous experience, which determines efficiency at time of start-up.” Instead, we would like to interpret it as saying that previous experience means that entrepreneurs are more confident about their ability and therefore adjustment costs do not cause them to be so prudent at start-up.

Therefore, one possibility to test our hypothesis against a theory of financing constraints, would be to test if in industries where adjustment costs are insignificant, age of entrepreneur does not have a significative effect on size. Alternatively, we could test whether the impact of age on size is correlated with some measure of the industry adjustment costs.

We could also attempt at explain the cross-country evidence that, even for narrowly defined industries, there are significant differences in the firm size distribution between countries. Bartelsman et. al. (2003, p. 26) interpret this as evidence that different institutional settings materialize themselves in different adjustment costs structures which then have implications for the firm size distributions:

...if certain administrative costs at entry are fixed, then the higher these costs (as in a number of European countries compared with the United Sates and the United Kingdom) the greater the disincentives for relatively small units to enter the market and then expand in the initial years. Likewise, post entry adjustments in employment may be hindered by tight hiring and firing restrictions and the latter are more restrictive in a number of European countries than in the United States.

Therefore, matching the structure of adjustment costs of each country, associated with their institutional settings, with the structure of their firm size distribution and its evolution for a given cohort of firms, would give a strong support for our hypothesis. Similarly, matching cross-industry differences in growth and firm size to cross-industry industry variation in adjustment costs, would also give a strong support to our theory. These are empirical projects under way.
A Appendix: Proofs

Proof of proposition 1.

First note that, when considering the best marginal adjustment, if it is optimal for the firm to adjust upwards, then we must solve

\[ A_H = [\alpha L_{t+\tau}^{\alpha-1} \theta_{t+\tau}^* - (w + H_p)] + \beta \frac{\partial E_{t+\tau}}{\partial L_{t+\tau}} \left\{ \max \left\{ V^O \left( L_{t+\tau} \right), V^S \left( L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 1 \right) \right\} \right\} = 0, \]

and if it is optimal for the firm to adjust downwards, we must solve

\[ A_F = [\alpha L_{t+\tau}^{\alpha-1} \theta_{t+\tau}^* - (w - F_p)] + \beta \frac{\partial E_{t+\tau}}{\partial L_{t+\tau}} \left\{ \max \left\{ V^O \left( L_{t+\tau} \right), V^S \left( L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 1 \right) \right\} \right\} = 0. \]

Now, the derivative can be rewritten as

\[
\frac{\partial E_{t+\tau}}{\partial L_{t+\tau}} \left\{ \max \left\{ V^O \left( L_{t+\tau} \right), V^S \left( L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 1 \right) \right\} \right\} = \int_{\Theta^O} \frac{\partial V^O \left( \cdot \right)}{\partial L_{t+\tau}} \, dF \left( \theta_{t+\tau+1}^* \mid \theta_{t+\tau}^*, \tau \right) + \int_{\Theta^{SD}} \frac{\partial V^{SD} \left( \cdot \right)}{\partial L_{t+\tau}} \, dF \left( \theta_{t+\tau+1}^* \mid \theta_{t+\tau}^*, \tau \right) + \int_{\Theta^{SN}} \frac{\partial V^{SN} \left( \cdot \right)}{\partial L_{t+\tau}} \, dF \left( \theta_{t+\tau+1}^* \mid \theta_{t+\tau}^*, \tau \right),
\]

where some of the regions might be empty, and while applying the Leibniz' rule we took into account the fact that at the frontiers the value functions assume the same value.

For each of the above derivatives we have

\[ \frac{\partial V^O \left( L_{t+\tau} \right)}{\partial L_{t+\tau}} = -O_p, \]

\[ \frac{\partial V^{SD} \left( L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 1 \right)}{\partial L_{t+\tau}} \bigg|_{\theta_{t+\tau+1}^* \in \Theta^{SD} \left( L_{t+\tau}, \tau + 1 \right)} = -F_p \]

\[ = [\alpha L_{t+\tau+1}^{\alpha-1} \theta_{t+\tau+1}^* - w] + \beta \frac{\partial E_{t+\tau+1}}{\partial L_{t+\tau+1}} \left\{ \max \left\{ V^O \left( L_{t+\tau+1} \right), V^S \left( L_{t+\tau+1}, \theta_{t+\tau+1}^*, \tau + 2 \right) \right\} \right\}, \]

\[ \frac{\partial V^{SN} \left( L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 1 \right)}{\partial L_{t+\tau}} = [\alpha L_{t+\tau+1}^{\alpha-1} \theta_{t+\tau+1}^* - w] + \beta \frac{\partial E_{t+\tau+1}}{\partial L_{t+\tau+1}} \left\{ \max \left\{ V^O \left( L_{t+\tau} \right), V^S \left( L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 2 \right) \right\} \right\}. \]
rewrites the above derivative as

\[
\frac{\partial V^{SU}(L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 1)}{\partial L_{t+\tau}} \bigg|_{\theta_{t+\tau+1}^* \in \Theta^{SU}(L_{t+\tau}, \tau + 1)} = H_p \\
= \left[ \alpha L_{t+\tau+1}^{\alpha-1} \theta_{t+\tau+1}^* - w \right] + \beta \frac{\partial E_{t+\tau+1} \left[ \max \left\{ V^O \left( L_{t+\tau+1} \right), V^S \left( L_{t+\tau+1}, \theta_{t+\tau+1}^*, \tau + 2 \right) \right\} \right]}{\partial L_{t+\tau+1}}
\]

where we have used the fact that \( A_H = 0 \), when it is optimal to adjust upwards, and \( A_F = 0 \), when it is optimal to adjust downwards. Therefore, we have

\[
\frac{\partial E_{t+\tau} \left[ \max \left\{ V^O \left( L_{t+\tau} \right), V^S \left( L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 1 \right) \right\} \right]}{\partial L_{t+\tau}}
\]

\[
= E_{t+\tau} \left( \chi_{t+\tau+1}^* \left( -O_p \right) + \left( 1 - \chi_{t+\tau+1}^* \right) \left\{ \alpha L_{t+\tau+1}^{\alpha-1} \theta_{t+\tau+1}^* - w \right\} \right.
\]

\[
+ \beta \frac{\partial E_{t+\tau+1} \left[ \max \left\{ V^O \left( L_{t+\tau+1}^* \right), V^S \left( L_{t+\tau+1}, \theta_{t+\tau+1}^*, \tau + 2 \right) \right\} \right]}{\partial L_{t+\tau+1}} \bigg|_{\theta_{t+\tau+1}^* \in \Theta^{SU}(L_{t+\tau+1}, \tau + 2)}
\]

Because we have a contraction mapping, using the law of iterated expectations, we can rewrite the above derivative as

\[
\frac{\partial E_{t+\tau} \left[ \max \left\{ V^O \left( L_{t+\tau} \right), V^S \left( L_{t+\tau}, \theta_{t+\tau+1}^*, \tau + 1 \right) \right\} \right]}{\partial L_{t+\tau}}
\]

\[
= \sum_{s=1}^{\infty} E_{t+\tau} \beta^{s-1} \left\{ \tilde{\chi}_{t+\tau+s}^* \left( -O_p \right) + \tilde{\chi}_{t+\tau+s}^* \left[ \alpha L_{t+\tau+s}^{\alpha-1} \theta_{t+\tau+s}^* - w \right] \right\}
\]

The result now follows by plugging the expression for the derivative in \( A_H \), and \( A_F \).

Finally, because there are fixed adjustment costs, adjustment will only occur if it proportionates a better outcome than non adjustment. \( \blacksquare \)

**Proof of corollary 1.**

First note that we can rewrite the LHS of (6) and (7) as follows

\[
MB \left( L_{t+\tau-1}, \theta_{t+\tau}^*, \tau \right) = \left( \alpha L_{t+\tau-1}^{\alpha-1} \theta_{t+\tau}^* - w \right) + \beta E_{t+\tau} \tilde{\chi}_{t+\tau+1}^* \left( -O_p \right)
\]

\[
+ \beta E_{t+\tau} \left\{ \tilde{\chi}_{t+\tau+1}^* \left( \alpha L_{t+\tau+1}^{\alpha-1} \theta_{t+\tau+1}^* - w \right) + \sum_{s=1}^{\infty} E_{t+\tau+1} \beta^s \left[ \tilde{\chi}_{t+\tau+1+s}^* \left( -O_p \right) \right.ight.
\]

\[
+ \tilde{\chi}_{t+\tau+1+s}^* \left( \alpha L_{t+\tau+1+s}^{\alpha-1} \theta_{t+\tau+1+s}^* - w \right] \left( \alpha L_{t+\tau+1+s}^{\alpha-1} \theta_{t+\tau+1+s}^* - w \right) \right\}.
\]

Taking into account that \( \tilde{\chi}_{t+\tau+1}^* \tilde{\chi}_{t+\tau+1+s}^* = \tilde{\chi}_{t+\tau+1+s}^* \tilde{\chi}_{t+\tau+1+s}^* \tilde{\chi}_{t+\tau+1+s}^* = \tilde{\chi}_{t+\tau+1+s}^* \tilde{\chi}_{t+\tau+1+s}^* \tilde{\chi}_{t+\tau+1+s}^* = 1 - \chi_{t+\tau+1}^* \), and \( \tilde{\chi}_{t+\tau+1}^* = \chi_{t+\tau+1}^* \), then we get the stated result. \( \blacksquare \)

**Proof of proposition 3.**

The result concerning the posterior distribution of \( \theta_{t+\tau+j} \) follows directly from

\[
\ln \left( \theta_{t+\tau+j} \right) \mid \Omega_{t+\tau} = \mu \mid \Omega_{t+\tau} + \varepsilon_{t+\tau+j}, \mu \mid \Omega_{t+\tau} \sim N \left( Y_{\tau}, S_{\tau} \right).
\]
For the distribution of $\theta^*_{t+\tau+j}$ conditional on $(\theta^*_{t+\tau}, \tau)$, we use the fact that

$$\ln (\theta^*_{t+\tau+j} \mid \Omega_{t+\tau}) = Y_{\tau+j} \mid \Omega_{t+\tau} + \frac{1}{2} (S_{\tau+j} + \sigma^2)$$

$$Y_{\tau+j} = \sigma^{-2} S_{\tau+j} \sum_{s=\tau}^{\tau+j-1} \eta_{t+s} + \frac{S_{\tau+j} Y_{\tau}}{S_{\tau}}$$

$$S_{\tau+j} = S_{\tau} - \sigma^{-2} S_{\tau+j} S_{\tau},$$

$$\eta_{t+s} \mid \Omega_{t+\tau} \sim N (Y_{\tau}, S_{\tau} + \sigma^2), \quad \text{Cov} (\eta_{t+s}, \eta_{t+s'} \mid \Omega_{t+\tau}) = \text{Var} (\mu \mid \Omega_{t+\tau}) = S_{\tau}, \ s, s' \geq \tau,$$

so that, in the end, we get

$$E [\ln (\theta^*_{t+\tau+j}) \mid \Omega_{t+\tau}] = Y_{\tau} + \frac{1}{2} (S_{\tau+j} + \sigma^2),$$

$$\text{Var} [\ln (\theta^*_{t+\tau+j}) \mid \Omega_{t+\tau}] = S_{\tau} - S_{\tau+j}.$$  

From here the result follows by noting that $\ln (\theta^*_{t+\tau}) = Y_{\tau} + \frac{1}{2} (S_{\tau} + \sigma^2)$. ■

**Proof of proposition 4.**

First of all, note that after period $t + T - 1$ the optimization problem is time invariant, since there is no uncertainty concerning $E (\theta)$. Therefore, for periods $t + T + s, s \geq 0$, the optimization problem becomes

$$V^S (\theta^*_{t+T}, L_{t+T+s-1}, T) = \max_{L_{t+T+s} \geq 0} \left\{ \left[ (\Pi (L_{t+T+s}, \theta^*_{t+T}) - C^S (L_{t+T+s}, L_{t+T+s-1})) + \beta \{ \chi_{t+T+s} [W - C^O (L_{t+T+s})] + (1 - \chi_{t+T+s}) V^S (\theta^*_{t+T}, L_{t+T+s}, T) \} \right] \right\}.$$  

Consider a firm that is in the industry at time $t + T + s - 1$, and assume that the firm has decided to remain in the industry at time $t + T + s$. We now prove that the employment decision of this firm will be the same in both periods. For this, the essential assumption is that

$$C^S (L_{t+T+s}, L^*_{t+T+s-1}) + C^S (L^*_{t+T+s-1}, L_{t+T+s-2}) \geq C^S (L_{t+T+s}, L_{t+T+s-2}),$$

where $L^*_{t+T+s-1} = L_{t+T+s-1} (\theta^*_{t+T}, L_{t+T+s-2}, T)$. That is, it is less costly to adjust in one step rather than in two. With this reasonable assumption, we have

$$\Pi (L_{t+T+s}, \theta^*_{t+T}) - C^S (L_{t+T+s}, L^*_{t+T+s-1}) + \beta V^S (\theta^*_{t+T}, L_{t+T+s})$$

$$= \Pi (L_{t+T+s}, \theta^*_{t+T}) - [C^S (L_{t+T+s}, L^*_{t+T+s-1}) + C^S (L^*_{t+T+s-1}, L_{t+T+s-2})]$$

$$+ \beta V^S (\theta^*_{t+T}, L_{t+T+s}) + C^S (L^*_{t+T+s-1}, L_{t+T+s-2})$$

$$\leq \Pi (L_{t+T+s}, \theta^*_{t+T}) - C^S (L_{t+T+s}, L_{t+T+s-2}) + \beta V^S (\theta^*_{t+T}, L_{t+T+s})$$

$$+ C^S (L^*_{t+T+s-1}, L_{t+T+s-2})$$

$$\leq V^S (\theta^*_{t+T}, L_{t+T+s-2}) + C^S (L^*_{t+T+s-1}, L_{t+T+s-2})$$

$$= \Pi (L^*_{t+T+s-1}, \theta^*_{t+T}) - C^S (L^*_{t+T+s-1}, L_{t+T+s-2}) + \beta V^S (\theta^*_{t+T}, L^*_{t+T+s-1})$$

$$+ C^S (L^*_{t+T+s-1}, L_{t+T+s-2})$$

$$= V^{SN} (\theta^*_{t+T}, L^*_{t+T+s-1}).$$
We conclude that at time $t + T + s$ it is optimal for the firm to set $L_{t+T+s}^* = L_{t+T+s-1}^*$.

We now prove that if the firm does not decide to exit at time $t + T + s$, then it will remain in the industry at time $t + T + s + 1$. Because the firm has decided to stay at time $t + T + s$, then $V^S (\theta_{t+T}, L_{t+T+s-1}^*, T) \geq W - C^O (L_{t+T+s-1}^*)$. But since next period the firm will not change its employment policy, that is, $L_{t+T+s}^* = L_{t+T+s-1}^*$, then

$$V^S (\theta_{t+T}, L_{t+T+s}^*; T) \geq V^S (\theta_{t+T}, L_{t+T+s-1}^*; T) \geq V^O (L_{t+T+s-1}^*) = V^O (L_{t+T+s}^*).$$

In conclusion, all adjustments will be made at time $t + T$, and after this period the firm will repeat the same decision. This means, that for optimization purposes we can solve the problem

$$V^S (\theta_{t+T}, L_{t+T-1}^*; T) = \max_{L_{t+T}} \left\{ \frac{1}{1 - \beta} \Pi (L_{t+T}; \theta_{t+T}) - C^S (L_{t+T}; L_{t+T-1}) \right\},$$

and determine the exit decision such that $V^S (\theta_{t+T}^*, L_{t+T-1}^*; T) \geq W - C^O (L_{t+T-1}^*)$.

**Proof of proposition 5.**

Note that

$$f (\theta_{t+s}^*, L_{t+s-1} \mid NE^{s-1}) = \frac{\int_{NE=1} f (\theta_{t+s}^*, L_{t+s-1}; \theta_{t+s-1}; L_{t+s-2} \mid NE^{s-2}) d\theta_{t+s-1} dL_{t+s-2}}{\int_{NE=1} f (\theta_{t+s-1}^*, L_{t+s-2} \mid NE^{s-2}) d\theta_{t+s-1} dL_{t+s-2}},$$

and the integrand in the numerator can be computed as

$$f (\theta_{t+s}^*, L_{t+s-1}; \theta_{t+s-1}; L_{t+s-2} \mid NE^{s-2})$$

$$= f (\theta_{t+s}^* \mid \theta_{t+s-1}^*, s - 1) f (L_{t+s-1}; \theta_{t+s-1}; L_{t+s-2} \mid NE^{s-2})$$

$$= f (\theta_{t+s}^* \mid \theta_{t+s-1}^*, s - 1) f (L_{t+s-1} \mid \theta_{t+s-1}^*, L_{t+s-2}, NE^{s-2})$$

$$\times f (L_{t+s-1}; \theta_{t+s-1}^*, L_{t+s-2} \mid NE^{s-2})$$

$$= f (\theta_{t+s}^* \mid \theta_{t+s-1}^*, s - 1) \mathbb{1} (L_{t+s-1} = L_{t+s-1} (\theta_{t+s-1}^*, L_{t+s-2}))$$

$$\times f (L_{t+s-1}; \theta_{t+s-1}^*, L_{t+s-2} \mid NE^{s-2}),$$

**B Appendix: Two-Period Model With Symmetric Adjustment Costs**

In this appendix, we show that in a two-period model with symmetric adjustment costs, there will be a tendency to start smaller and grow over time when the distribution of the productivity parameter is skewed to the right and the degree of decreasing returns to scale, $\alpha$, is higher than $1/2$. We analyze separately the cases of fixed and proportional adjustment costs, and do not consider the exit decision. By doing so it would only complicate the analysis, without making it clear the effect of these two factors.
B.1 Fixed Adjustment Costs

We will consider a simplified version of the general model we have analyzed in the text. In this version, the firm is only allowed to remain in the industry for two periods, and it learns completely its productivity after the first period production, so that in the second period the firm faces no uncertainty. In this subsection, we also assume that the firm faces only a symmetric fixed adjustment cost: $K_f$. Therefore, the optimization problem is given by

$$V_1 \equiv \max_{L_1, L_2} E \{ \Pi (L_1, \theta) + \beta [\Pi (L_2, \theta) - 1 (L_2 \neq L_1) K_f] \}$$

$$\equiv \max_{L_1} E \left\{ \Pi (L_1, \theta) + \beta \max_{L_2} [\Pi (L_2, \theta) - 1 (L_2 \neq L_1) K_f] \right\},$$

Solving the model by backwards induction, if the firm adjusted in the second period, the optimal employment and the profit under adjustment would have been,

$$L_2 = \bar{L} (\theta) = \left( \frac{\alpha \theta}{w} \right)^{\frac{1}{1-\alpha}},$$

$$V^A (\theta) \equiv (1 - \alpha) \left( \frac{\alpha}{w} \right)^{\frac{\alpha}{1-\alpha}} \theta^{\frac{1}{1-\alpha}} - K.$$

If the firm did not adjust, we would have the same employment as in the first period

$$L_2 = L_1,$$

$$V^{NA} (L_1, \theta) \equiv L_1^\alpha \theta - wL_1.$$

Because $V (L_1, \theta) = \max \{ V^A (\theta), V^{NA} (L_1, \theta) \}$, then, conditional on $L_1$, the firm will not adjust only if

$$V^{NA} (L_1, \theta) \geq V^A (\theta) \Leftrightarrow \theta \in [\theta_1^{NA} (L_1), \theta_2^{NA} (L_1)]$$

where $\theta_2^{NA}$ and $\theta_1^{NA}$ are the higher and the lower roots of $V^{NA} (L_1, \theta) - V^A (\theta) = 0$. If the lower root of this equation is negative, then we make $\theta_1^{NA} = 0$. For future reference, we can also identify the point at which the slope of $V^{NA} (L_1, \theta)$ equals the slope of $V^A (\theta)$:

$$\theta^{BN} (L_1) \equiv L_1^{1-\alpha \frac{w}{\alpha}}. \text{ This is the value of } \theta \text{ for which the first period employment level would maximize the second period profits. Therefore, the employment decision and the value function in the second period are defined by}$$

$$L_2^* (L_1, \theta) = \begin{cases} \bar{L} (\theta), & \text{if } \theta \notin [\theta_1^{NA} (L_1), \theta_2^{NA} (L_1)] \smallskip L_1, & \text{if } \theta \in [\theta_1^{NA} (L_1), \theta_2^{NA} (L_1)] \end{cases},$$

$$V (L_1, \theta) = \begin{cases} V^A (\theta), & \text{if } \theta \notin [\theta_1^{NA} (L_1), \theta_2^{NA} (L_1)] \smallskip V^{NA} (L_1, \theta), & \text{if } \theta \in [\theta_1^{NA} (L_1), \theta_2^{NA} (L_1)] \end{cases}.$$
with the general result that in discrete and finite time the optimal bands are state dependent. As a particular case of proposition 1, the optimality condition for first period employment is

\[ \alpha L_1^{\alpha-1} E(\theta) - w + \beta \int_{\theta_1^{NA}(L_1)}^{\theta_2^{NA}(L_1)} \alpha L_1^{\alpha-1} \theta - w \ dF(\theta) = 0. \]  

(12)

It is obvious that, in general, this equation can only be solved numerically for \( L_1 \). However, we can derive some alternative analytic expressions

\[ L_1^* = \left( \frac{\alpha E(\theta)}{w - A(L_1)} \right)^\frac{1}{\alpha}, \]

\[ A(L_1) = \beta \Pr[\theta_1^{NA} \leq \theta \leq \theta_2^{NA}] \{ \alpha L_1^{\alpha-1} E[\theta \mid \theta_1^{NA} \leq \theta \leq \theta_2^{NA}] - w \}, \]

\[ L_1^* = \frac{\alpha E(\theta)}{w} \times \left( \frac{1 + \beta \Pr[\theta_1^{NA} \leq \theta \leq \theta_2^{NA}] E(\theta \mid \theta_1^{NA} \leq \theta \leq \theta_2^{NA})}{1 + \beta \Pr[\theta_1^{NA} \leq \theta \leq \theta_2^{NA}]} \right)^\frac{1}{\alpha}. \]

From these expressions, we conclude that

\[ L_1^* < \bar{L}[E(\theta)] \iff E[\theta \mid \theta_1^{NA} \leq \theta \leq \theta_2^{NA}] < E(\theta) = \bar{\theta} \{ \bar{L}[E(\theta)] \} \]

\[ \iff \alpha L_1^{*(\alpha-1)} E[\theta \mid \theta_1^{NA} \leq \theta \leq \theta_2^{NA}] - w < 0. \]

That is, firms will start smaller if doing so implies, conditional on nonadjustment, an expected value for labor higher than optimal, which is equivalent to say that starting smaller implies an expected productivity, conditional on nonadjustment, smaller than the unconditional expected productivity. The conditions under which this will occur seem to be related to the properties of the production function and the properties of the distribution for \( \theta \). We analyze this in the following proposition.

**Proposition 6** Under some regularity conditions for \( F(\theta) \), in the two-period model under consideration with symmetric fixed adjustment costs, if \( \alpha > \frac{1}{2} \) and if \( F(\theta) \) is a unimodal skewed to the right distribution function, then \( L_1^* < \bar{L}[E(\theta)] \).

**Proof.**

We first prove that if \( F(\theta) \) is symmetric and \( \alpha > \frac{1}{2} \) then \( L_1^* < \bar{L}[E(\theta)] \). Note that, given the definitions of \( \theta_1^{NA} \), \( \theta_2^{BN} \), and \( \theta_2^{NA} \) and assuming that \( \theta_1^{NA} \{ \bar{L}[E(\theta)] \} > 0 \), we can prove that\(^{15}\)

\[ \alpha > \frac{1}{2} \iff \theta_2^{BN}(L_1) > \frac{\theta_1^{NA}(L_1) + \theta_2^{NA}(L_1)}{2}. \]

Consider now the case where \( L_1 = \bar{L}[E(\theta)] \). Then \( \alpha \bar{L}[E(\theta)]^{\alpha-1} E(\theta) - w = 0 \). Also, because \( f(\theta) \) is symmetric and \( \alpha > \frac{1}{2} \), then

\[ \theta_2^{BN} \{ \bar{L}[E(\theta)] \} = E(\theta) > \frac{\theta_1^{NA} \{ \bar{L}[E(\theta)] \} + \theta_2^{NA} \{ \bar{L}[E(\theta)] \}}{2}, \]

\(^{15}\) Just note that, for \( \alpha = \frac{1}{2} \), \( V^A(\theta) - V^{NA}(L_1, \theta) \) is a parabola, in which case \( \theta_2^{BN}(L_1) = \frac{\theta_1^{NA}(L_1) + \theta_2^{NA}(L_1)}{2} \).
implies

\[
E(\theta) > E[\theta | \theta_1^{NA} \{ \bar{L}[E(\theta)] \}] \leq \theta \leq \theta_2^{NA} \{ \bar{L}[E(\theta)] \}
\]

Therefore, if \( \alpha > \frac{1}{2} \), the FOC (12) will be negative, and if an interior maximum exists, with the objective function being concave in the interval \([L_1^*, \bar{L}[E(\theta)]\), then the solution will be characterized by \( L_1^* < \bar{L}[E(\theta)] \). This result is important per se since in the absence of adjustment costs, it is always optimal to set \( L_1^* = \bar{L}[E(\theta)] \).

Now we assume \( \alpha \) equal to \( \frac{1}{2} \) and \( F(\theta) \) skewed to the right (e.g., the lognormal distribution). In this case, even with \( \alpha = \frac{1}{2} \), on “average”, values smaller than \( \theta^{BN} \) receive more weight than values greater than \( \theta^{BN} \) so that

\[
E[\theta | \theta_1^{NA} \{ \bar{L}[E(\theta)] \}] \leq \theta \leq \frac{\theta_1^{NA} \{ \bar{L}[E(\theta)] \} + \theta_2^{NA} \{ \bar{L}[E(\theta)] \}}{2} = \theta^{BN} \{ \bar{L}[E(\theta)] \} = E(\theta).
\]

This implies, as before, that the FOC will be negative and, with concavity, \( L_1^* < \bar{L}[E(\theta)] \).

### B.2 Proportional Adjustment Costs

In what follows, we analyze the two period model when there is a symmetric proportional adjustment cost of \( K_p \). The optimization problem is now given by

\[
V_I \equiv \max_{L_1, L_2} E\{ \Pi(L_1, \theta) + \beta [\Pi(L_2, \theta) - 1 (L_2 \neq L_1) | L_2 - L_1 | K_p] \}
\]

\[
\equiv \max_{L_1} E\{ \Pi(L_1, \theta) + \beta \max_{L_2} [\Pi(L_2, \theta) - 1 (L_2 \neq L_1) | L_2 - L_1 | K_p] \}.
\]

By solving the second period problem, we get

\[
L_2^*(L_1, \theta) = \begin{cases} 
\bar{L}^U(\theta) = \left( \frac{w - K_p}{\alpha w + \alpha K_p} \right)^{1 - \alpha}, & \text{if } \theta > \theta_2^{NA}(L_1, K_p), \\
L_1, & \text{if } \theta \in [\theta_1^{NA}(L_1, K_p), \theta_2^{NA}(L_1, K_p)], \\
\bar{L}^D(\theta) = \left( \frac{w - K_p}{\alpha w - K_p} \right)^{1 - \alpha}, & \text{if } \theta < \theta_1^{NA}(L_1, K_p),
\end{cases}
\]

where

\[
\theta_1^{NA}(L_1, K_p) = \frac{w - K_p}{\alpha L_1^{\alpha - 1}}, \quad \theta_2^{NA}(L_1, K_p) = \frac{w + K_p}{\alpha L_1^{\alpha - 1}}. \tag{14}
\]

Because \( V(L_1, \theta) = \max \{ V^{AD}(L_1, \theta), V^{NA}(L_1, \theta), V^{AU}(L_1, \theta) \} \), then we also have

\[
V(L_1, \theta) = \begin{cases} 
V^{AU}(L_1, \theta) = (1 - \alpha) \left( \frac{\alpha}{w + K_p} \right)^{1 - \alpha} \theta^{1 - \alpha} + K_p L_1, & \text{if } \theta > \theta_2^{NA}(L_1, \theta), \\
V^{NA}(L_1, \theta) = L_2^* \theta - w L_1, & \text{if } \theta \in [\theta_1^{NA}(L_1, \theta), \theta_2^{NA}(L_1, \theta)], \\
V^{AD}(L_1, \theta) = (1 - \alpha) \left( \frac{\alpha}{w - K_p} \right)^{1 - \alpha} \theta^{1 - \alpha} - K_p L_1, & \text{if } \theta < \theta_1^{NA}(L_1, \theta).
\end{cases}
\]

28
As a particular case of proposition 1, the optimal level for first period employment will be determined from

\[
[\alpha L_1^{\alpha-1}E(\theta) - w] + \beta \left\{ (-K_p) \int_{\nu_1}^{\theta_1 N_A(L_1)} dF(\theta) + K_p \int_{\nu_1}^{\nu_2} dF(\theta) \right. \\
+ \left. \int_{\theta_1 N_A(L_1)}^{\theta_2 N_A(L_1)} [\alpha L_1^{\alpha-1}\theta - w] dF(\theta) \right\} = 0, \tag{15}
\]

It is obvious that, in general, this equation can only be solved numerically for \( L_1 \). In the following proposition we present one sufficient condition for \( L_1^* < \bar{L}[E(\theta)] \).

**Proposition 7** Under some regularity conditions for \( F(\theta) \), in the two-period model under consideration with symmetric proportional adjustment costs, if \( F(\theta) \) is a unimodal skewed to the right distribution function, then \( L_1^* < \bar{L}[E(\theta)] \).

**Proof.**

Assume that \( L_1 = \bar{L}[E(\theta)] \). Then, the LHS of (15) becomes

\[
FOC_p(K_p) = \int_{\nu_1}^{\theta_1 N_A(\bar{L}[E(\theta)])} (-K_p) dF(\theta) + \int_{\theta_1 N_A(\bar{L}[E(\theta)])}^{\nu_2} K_p dF(\theta) \\
+ \int_{\theta_1 N_A(\bar{L}[E(\theta)])}^{\theta_2 N_A(\bar{L}[E(\theta)])} [\alpha \bar{L}[E(\theta)]^{\alpha-1}\theta - w] dF(\theta).
\]

If we take a derivative of this expression with respect to \( K_p \), making use of the fact that \( \alpha L_1^{\alpha-1}\theta_1 N_A - w = -K_p \), and \( \alpha L_1^{\alpha-1}\theta_2 N_A - w = K_p \), then we get \( F \left( \theta_1 N_A \right) + [1 - F \left( \theta_2 N_A \right)] \). Because we are assuming that \( \theta_1 N_A \left\{ \bar{L}[E(\theta)] \right\} > 0 \), then if \( F(\theta) \) is smooth enough in the sense that there is only one \( K_p \) such that \( -F \left( \theta_1 N_A \left\{ \bar{L}[E(\theta)] \right\} \right) + [1 - F \left( \theta_1 N_A \left\{ \bar{L}[E(\theta)] \right\} \right)] = 0 \), then the above function will be have two possible maximums: one at \( K_p = 0 \), at which point is zero, and the other for \( K_p = K_{p,1} \), where \( \theta_1 N_A \left\{ \bar{L}[E(\theta)] \right\} = \nu_1 \). We analyze the second case. From this, we have for the above expression \( FOC_p \leq \max \{0, FOC_p \left( K_p = K_{p,1} \right) \} \)

\[
FOC_p(K_p = K_{p,1}) = \int_{\theta_1 N_A(\bar{L}[E(\theta)])}^{\nu_2} K_{p,1} dF(\theta) + \int_{\nu_1}^{\theta_2 N_A(\bar{L}[E(\theta)])} \left[ \alpha \bar{L}[E(\theta)]^{\alpha-1}\theta - w \right] dF(\theta) \\
< \int_{\nu_1}^{\nu_2} \left[ \alpha \bar{L}[E(\theta)]^{\alpha-1}\theta - w \right] dF(\theta) = 0.
\]

We then conclude that for \( K_p > 0 \), we have \( FOC_p < 0 \). Because we can prove that the profit function is concave, that is, the derivative of (15) with respect to \( L_1 \) is

\[
\alpha (\alpha - 1) L_1^{\alpha-2} [E(\theta) + \beta E \left( \theta \mid \theta_1 N_A \leq \theta \leq \theta_2 N_A \right)] < 0,
\]

then \( L_1^* < \bar{L}[(\theta)] \).
Appendix: Computational Algorithm

We present in various steps the algorithm employed to solve finite learning horizon model.

(i) Discretization of $\theta$

Because \textit{ex ante} $\theta_{t+T}^{*} = \exp\left(\mu + \frac{1}{2}\sigma^2\right) \sim \log N\left(\mu + \frac{1}{2}\sigma^2, \sigma^2_\mu\right)$, we have followed Jérome Adda and Russell Cooper (2003), adapting it to the lognormal distribution, so that if we let $\{\bar{\theta}_j\}_{j=1}^{N_\theta+1}$ be the limits of the intervals that discretize the real line, and $\{\theta_j\}_{j=1}^{N_\theta}$ be the mass points, then

$$
\bar{\theta}_j = \exp\left\{\bar{\mu} + \frac{1}{2}\sigma^2 + \sigma_\mu \Phi^{-1}\left(\frac{j-1}{N}\right)\right\}
$$

$$
\theta_j = N_\theta \exp\left(\bar{\mu} + \frac{1}{2}(\sigma^2 + \sigma^2_\mu)\right) \times
$$

$$
\left[\Phi\left(\frac{\ln(\bar{\theta}_{j+1}) - (\bar{\mu} + \frac{1}{2}\sigma^2 + \sigma^2_\mu)}{\sigma_\mu}\right) - \Phi\left(\frac{\ln(\bar{\theta}_j) - (\bar{\mu} + \frac{1}{2}\sigma^2 + \sigma^2_\mu)}{\sigma_\mu}\right)\right],
$$

and we have used $N_\theta = 20$ points.

(ii) Transition probability matrices

Once again, we have followed Adda and Cooper

$$
\pi_{ij}^s = \Pr\left\{\theta_{t+s}^* \in [\theta_j, \bar{\theta}_{j+1}] \mid \theta_{t+s-1}^* \in [\bar{\theta}_i, \theta_{i+1}]\right\}
$$

$$
= \left[\Phi\left(\frac{\ln(\bar{\theta}_{t+1}) - (\bar{\mu} + \frac{1}{2}(S_{s-1} + \sigma^2))}{\sqrt{\sigma^2_\mu - S_{s-1}}}\right) - \Phi\left(\frac{\ln(\bar{\theta}_t) - (\bar{\mu} + \frac{1}{2}(S_{s-1} + \sigma^2))}{\sqrt{\sigma^2_\mu - S_{s-1}}}\right)\right]^{-1}
$$

$$
\times \int^{\ln(\bar{\theta}_{t+1})}_{\ln(\bar{\theta}_t)} \left[\Phi\left(\frac{\ln(\theta_{t+1}) - (\ln(\theta_{t+s-1}) - \frac{1}{2}(S_{s-1} - S_s))}{\sigma_\mu}\right) - \Phi\left(\frac{\ln(\theta_j) - (\ln(\theta_{t+s-1}) - \frac{1}{2}(S_{s-1} - S_s))}{\sigma_\mu}\right)\right] d\ln(\theta_{t+s-1}),
$$

where $\ln(\theta_{t+s-1}) = Y_{s-1} + \frac{1}{2}(S_{s-1} + \sigma^2) \sim N\left(\bar{\mu} + \frac{1}{2}(S_{s-1} + \sigma^2), \sigma^2_\mu - S_{s-1}\right)$.

(iii) Discretization of $L$

We have used the decision rules for problem (11) in case of hiring and in case of firing, from which we considered

$$
\ln(L) \sim N\left(\mu_L, \sigma^2_L\right),
$$

$$
\mu_L = \frac{1}{1-\alpha}\left\{\bar{\mu} + \frac{1}{2}\sigma^2 + \ln\left(\frac{\alpha}{\sqrt{w + (1-\beta)H_p}[w - (1-\beta)F_p]}\right)\right\},
$$

$$
\sigma^2_L = \frac{1}{(1-\alpha)^2}\sigma^2_\mu,
$$

and used the same procedure to discretize $L$ as the one used for $\theta$, considering $N_L = 200$. 
(iv) Choice for $T$
We have chosen $T = 25$, so that $S_{T-1} < 0.004$.

(v) Updating rule for $p$
Given that in problem (8) the value function, when $p$ is included, is given by

$$
\frac{(1 - \alpha)}{(1 - \beta)}\left(\frac{\alpha}{w + (1 - \beta)I_p}\right)^{\frac{1}{\alpha}}(p\theta)^{\frac{1}{1-\alpha}} - I_f,
$$

then we consider

$$
p_{i+1} = p_i \left(\frac{W + I_f}{V' + I_f}\right)^{(1-\alpha)}.
$$

References


Zellner, Arnold (1971), *An Introduction to Bayesian Inference in Econometrics*, New York: John Wiley & Sons
\[ \alpha = 0.66, \beta = 0.95, w = 10, W = 100, E(\mu) = 2.5, \sigma_\mu = 0.26458, \sigma = 0.31623 \]

Case 1: \[ H_f = 0, H_p = 0, F_f = 0, F_p = 0, I_f = 50, I_p = 0, O_f = 0, O_p = 0, p = 1.1839 \]

Case 2: \[ H_f = 0, H_p = 0, F_f = 0, F_p = 0, I_f = 50, I_p = 0, O_f = 80, O_p = 0, p = 1.2277 \]
$\alpha = 0.66, \beta = 0.95, w = 10, W = 100, E(\mu) = 2.5, \sigma_{\mu} = 0.26458, \sigma = 0.31623$

Case 1: $H_f = 0, H_p = 0, F_f = 0, F_p = 0, I_f = 50, I_p = 0, O_f = 0, O_p = 0, p = 1.1839$

Case 2: $H_f = 0, H_p = 0, F_f = 5, F_p = 3, I_f = 50, I_p = 0, O_f = 5, O_p = 3, p = 1.2006$
\(\alpha = 0.66, \beta = 0.95, w=10, W=100, E(\mu) = 2.5, \sigma = 0.26458, \sigma = 0.31623\)

**Case 1:**
- \(H_f=0, F_f=0, I_f=50, O_f=80\), \(p=1.2277\)

**Case 2:**
- \(H_f=0, F_f=5, I_f=50, O_f=85\), \(p=1.2376\)
\( \alpha = 0.66, \beta = 0.95, w = 10, W = 100, E(\mu) = 2.5, \sigma_\mu = 0.26458, \sigma = 0.31623 \)

Case 1: \( H_f = 0, H_p = 0, F_f = 0, F_p = 0, I_f = 50, I_p = 0, O_f = 0, O_p = 0, p = 1.1839 \)

Case 2: \( H_f = 5, H_p = 3, F_f = 0, F_p = 0, I_f = 55, I_p = 3, O_f = 0, O_p = 0, p = 1.2303 \)
Figure 5

Exit Threshold for Periods $t+2$ and $t+\tau$

Average Employment Conditional on No Exit for Periods $t$ to $t+\tau$

Average Growth Rate of Employment for Periods $t+1$ to $t+\tau$

Proportion of Exiting Firms for Periods $t+1$ to $t+\tau$

$\alpha=0.66, \beta=0.95, w=10, W=100, E(\mu)=2.5, \sigma_\mu=0.26458, \sigma=0.31623$

Case 1: $H_f=0, H_p=0, F_f=0, F_p=0, l_f=50, l_p=0, Q_f=0, Q_p=0, p=1.1839$

Case 2: $H_f=2, H_p=1, F_f=5, F_p=3, l_f=52, l_p=1, Q_f=5, Q_p=3, p=1.2181$
Figure 6

- Exit Threshold for Periods $t+2$ and $t+\tau$
- Average Employment Conditional on No Exit for Periods $t$ to $t+\tau$
- Average Growth Rate of Employment for Periods $t+1$ to $t+\tau$
- Proportion of Exiting Firms for Periods $t+1$ to $t+\tau$

Parameters:
- $\alpha = 0.66, \beta = 0.95, w = 10, W = 100, E(\mu) = 2.5, \sigma_{t+2} = 0.26458, \sigma = 0.31623$
- Case 1: $H_f = 0, H_p = 0, F_f = 0, F_p = 0, I_f = 50, I_p = 0, O_f = 80, O_p = 0, p = 1.2277$
- Case 2: $H_f = 2, H_p = 1, F_f = 5, F_p = 3, I_f = 52, I_p = 1, O_f = 85, O_p = 3, p = 1.2531$
Figure 7

Exit Threshold for Periods t+2 and t+τ

Average Employment Conditional on No Exit for Periods t to t+τ

Average Growth Rate of Employment for Periods t+1 to t+τ

Proportion of Exiting Firms for Periods t+1 to t+τ

\[ \alpha = 0.66, \beta = 0.95, w = 10, W = 100, E(\mu) = 2.5, \sigma_\mu = 0.26458, \sigma = 0.31623 \]

Case 1: \( H_f = 0, H_p = 0, F_f = 0, F_p = 0, I_f = 50, I_p = 0, O_f = 0, O_p = 0, \) \( p = 1.1839 \)

Case 2: \( H_f = 5, H_p = 3, F_f = 2, F_p = 1, I_f = 55, I_p = 3, O_f = 2, O_p = 1, \) \( p = 1.2359 \)