NORGES HANDELSHØYSKOLE
Bergen, 17.06.2009

Estimating Interrelated Nonconvex Adjustment Costs

Magne Krogstad Asphjell

Supervisor: Professor Øivind Anti Nilsen

Master Thesis, Department of Economics

NORWEGIAN SCHOOL OF ECONOMICS

AND BUSINESS ADMINISTRATION

Summary

Adjustment costs associated with firms' acquirement or disposal of factors of production can make the individual firm unresponsive to changes in their environment. This is the reason why costs of adjustments is assumed to be one of the reasons why we observe firm and plant level adjustment patterns as unevenly distributed over time. Understanding what characterizes such costs is important for understanding firm level behaviour, but can also be important to understand dynamics of both capital and labor demand at macro level. Because several studies on adjustment costs and factor demand indicate correlation in the demand for production factors, this thesis aims to present a way to simultaneously estimate adjustment cost functions of two factors, including both convex and nonconvex components. Joint estimation also allows for possible existence of interrelations in the adjustment cost function. A likelihood function for estimation by Maximum Likelihood is derived, and results after estimation on simulated panel data is presented. The text shows how a likelihood function can be written to estimate adjustment cost parameters that can be traced directly back to a theoretical framework for adjustment costs and factor demand. It is also shown that under certain conditions, the procedure inhabits weaknesses regarding identification of model parameters, which should be improved for increased robustness.

Table of Contents

1. Introduction

When a producer of any non-storable good decides to change the level of output from one period to the next, the firm will have to make an investment in one or more production factors. The investment decision will, among other things, depend on expectations about future demand and access to new technology. Even though one would expect investment patterns to be relatively smooth over time, given that shocks in the environment are relatively small and occur relatively often, we most often observe firm level investment patterns as lumpy rather than smooth both for capital and labor (Bloom, 2009).

A typical explanation for this phenomenon is that what we call adjustment costs that arise when firms make investments, together with uncertainty about future shocks can make firms hesitant to making adjustments even if they face changes in the environment that affect their optimal levels of inputs. A gap between optimal and actual input levels would in particular arise if adjustment cost functions would contain some form of nonconvexities (Hamermesh and Pfann, 1996).

In empirical work on the role of adjustment costs, underlying theoretical models are often based on a one-factor analysis with one quasi-fixed production factor. At the same time, research has shown that results indicating quasi-fixity of one factor in reality can be caused by some form of interrelation in multi-factor demand. By interrelation we simply mean the possibility of the demand for one input affecting the demand for another, and vice versa.

The motivation behind this thesis is first to understand what characterizes firms' investment behavior in both capital and labor, and how this behavior is affected by adjustment costs of different characteristics. In addition, the role of possible interrelations in the cost functions in a two-factor setting will be investigated. Section 2 will give an overview of the theoretical foundation for discussing adjustment costs and the dynamic optimization problem firms face prior to changes in input factors. Section 3 introduces the concept of interrelations in adjustment cost functions, and presents a theoretical model to incorporate effects of this sort.

Next, the ambition is to show how a Maximum Likelihood approach can be applied to estimate parameters of adjustment cost functions that are both non-convex and possibly interrelated for two factors of production. The derivation of a suitable likelihood function with properties similar to a bivariate ordered probit model is shown in section 4. Estimation by the model derived is carried out using a simulated dataset containing moments often observed in real data. Basic descriptions of the data are presented in section 5. Finally, sections 6 and 7 provide summarizations of results and concluding remarks.

2. Adjustment Costs and Changes in Input Factors

2.1. What do we mean by investment behavior?

This thesis aims to investigate what characterizes the way that firms change the level of their production factors. In a simplified world, we would expect firms to make investments according to changes in input factor profitability. Based on a standard marginal value equilibrium condition we would expect that if the marginal value or the purchase price of one input factor unexpectedly rises or falls, we would see a corresponding positive or negative factor adjustment. This would certainly always be the case if there were no costs associated with investments (Abel and Eberly, 1994).

However, this is rarely the case in the real world, and research has shown that changes in factor levels are not necessarily instantaneous, especially when the firm is facing only small changes in prices or productivity. In addition we often observe investments to be relatively large in those periods in which they occur. When changes in input factors are unevenly distributed over time resulting in adjustment spikes, we say that investments are lumpy. In the opposite case, when changes in factor levels are evenly distributed over time and spikes are rarely observed, we say that investments are smooth.

At macro level, most data suggest that factor demand adjusts smoothly. However, with an increasing access to firm or plant level panel data, researchers have observed that investments in both capital and labor are made in a lumpy fashion. To explain the observed firm level adjustment patterns, researchers have throughout the last decades tried to study the possible role of nonconvex adjustment costs together with uncertainty about future shocks. The lasting high level of European unemployment is one of several economic issues where adjustment costs might play an important role (Hamermesh, 1997).

The goal of this analysis is to derive a simultaneous econometric model of firm-level investment behavior for two production inputs. To establish an understanding of what should be included in a model describing firms’ factor demand, we should have an overview of how the research on the subject has evolved, and some of the key results that have eventually lead to the derivation of models used today. Research has roughly developed from one-factor analyses of factor demand patterns, to investigation of the role of adjustment costs and further into the role of such costs in multi-factor models.

2.2. Early Findings on Factor Demand

In the early work by Jorgenson (1963), the author uses neoclassical theory of optimal accumulation of capital to present a theory of investment behavior. The author finds that because of lagged response, firms' demand for investments is not equal to their demand for capital. The work in Jorgenson's paper is done using aggregated data, and even though the findings are in line with what more recent research has tried to explain, Jorgenson states nothing about what might cause the slow response of capital investment. However, the econometric model presented uses the firm's maximization of discounted future profits as a way of explaining investment behavior, an approach still used in today's research.

Nadiri and Rosen (1969) investigate the degree of fixity of four different factors of production. The paper presents a two-stage model where the firm, in the first stage minimizes total long-run costs, and in the second stage decides optimal adjustments of production factors by minimizing discounted costs. In the model presented, production factors are defined as capital and labor in addition to utilization rates for both inputs. After an empirical investigation, the paper concludes that the utilization rates of capital and labor (in this order) are adjusted most rapidly, while the levels of capital and labor (in this order) are the least flexible of the four inputs.

Like Jorgensen, Nadiri and Rosen do not state anything about the role of possible adjustment costs when investment strategies are chosen, but the findings regarding investment dynamics are still interesting, and should be considered when comparing empirical findings in more recent studies. For example, the results indicate that the level of capital is relatively fixed compared to the other inputs. The authors' own interpretation of the results seems to be that the role of utilization rates, and to some extent the level of employment, is to maintain the level of production at a desired level, while the level of capital input is adjusted slowly. Although the explanation of the behavior is different, the findings regarding dynamics are similar to conclusions in more recent research.

In a generalized form of the Q-model[1] framework Galeotti and Schiantarelli (1991) treat labor as a quasi-fixed input factor while considering the level of capital as fixed when the adjustment decision regarding labor is taken. The paper presents estimates for aggregate data on the U.S. manufacturing sector, and the authors focus their analysis on the change of number of workers, not on the change of work hours, as this is consistent with a priori beliefs and earlier empirical findings. The results indicate a support for treating labor in addition to capital as quasi-fixed and possibly also allowing for interrelations between the two adjustment processes.

2.3. Costs of Adjustment

Judging from the available literature on the subject, it is fairly obvious that investment decisions in some way are interfered so that lumpy patterns can arise. To explain the slow response in firms' investments, theoretical models of factor adjustment have in the last decades been equipped to consider the existence of adjustment costs. If there were no costs associated with changing the level of production inputs, firms would certainly react instantly to any change in the environment. Introducing adjustment costs in our investment models shows how the adjustment decision becomes a problem that is among several things affected by expectations about future events. In other words, we recognize how adjustment costs can play a key role in explaining observed behavior. But what then is the source of such costs, and how can they best be characterized? Different assumptions about functional form have different economic interpretations and give different predictions about expected firm behavior. This has indeed been the focus of much research. Hamermesh and Pfann (1996) give us an overview of the topic.

[1] See Tobin, J. (1969) for details on Tobin's q. See for example Blundell et. al. (1992) for an application to capital investment.

2.4. Functional Form

A key question when discussing qualities of adjustment costs is their possible shape, and which functional form we can assume for the best possible approximation of real life conditions. Before making any assumptions in this thesis it would be sensible to include a brief discussion on what we can assume regarding the shape of so-called adjustment costs, and what implications we make using different assumptions about functional form. Hamermesh and Pfann (1996) describe four basic examples of modeling the structure of costs of adjustment. The following section is based on their work, and will illustrate different possible ways of specifying adjustment costs and their implications.

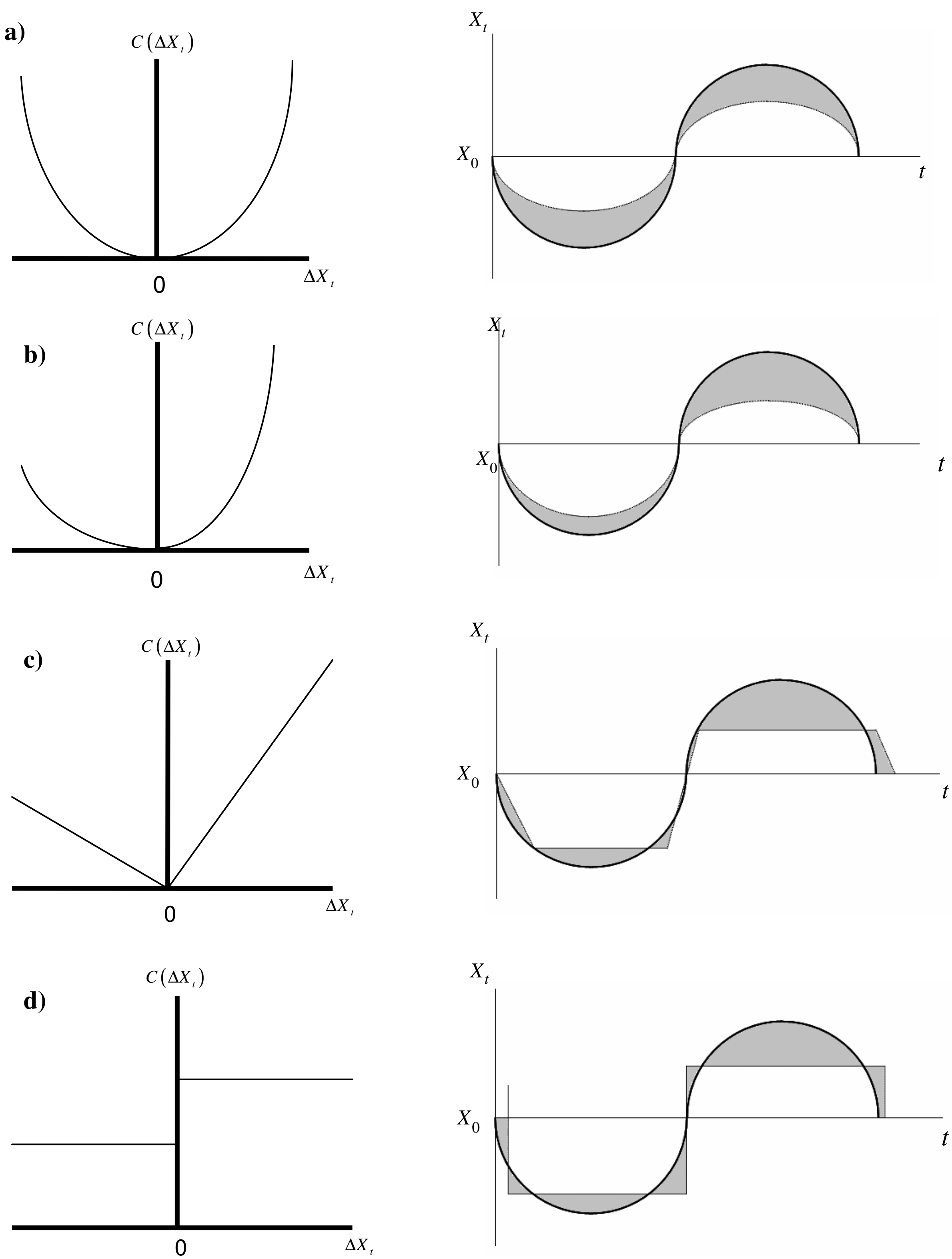
2.4. a) Symmetric convex adjustment costs

Much of the early literature on the subject assumes that the costs of adjustments are convex and symmetric around $\Delta X = 0$. This is in particular true about the research that is based on the q-theory of investment. We can write a general example of a symmetric and convex specification as

$$C(\Delta X) = \frac{1}{2} b \left[\Delta X \right]^2, b > 0$$

and the function is shown in figure (2.1.a). From the figure, we clearly see that adjustments under this specification are made continuously, although there is a clear gap between optimal input levels and the desired adjustments. A specification of this kind imply that we assume the costs of expanding or contracting the levels of capital or labor force to be equal, which is hardly a reasonable assumption in most cases. The popularity is no doubt caused almost only by analytical tractability. As Hamermesh and Pfann note: "Simply imposing (…), no matter how many times it has been done, in no way speaks to the correctness of the underlying assumption." In other words, one is not advised to assume the simplest possible form, just because many have done so in earlier research.

Figure 2.1: Adjustment cost functions



2.4. b) Asymmetric convex adjustment costs

Because there is no reason why the marginal cost of increasing the level of one production input would be the same as that of an equal size decrease, Hamermesh and Pfann go on to describe a second possible approximation as Asymmetric Convex Costs. One particular function of this type can be written as:

$$C(\Delta X) = .5b[\Delta X]^2 - c\Delta X + \exp(c\Delta X) - 1$$

and a graphical representation is found in figure (2.1.b). Again we see that adjustments are made continuously, although the gaps between optimal and actual input levels now differ in accordance with the underlying asymmetry.

2.4. c) Piecewise linear adjustment costs

A third functional form also opens for possible asymmetry. Piecewise linear costs though, as the name suggests, assumes linear adjustment costs which are proportional to changes in the production factor in question. Hamermesh and Pfann writes this cost function as:

$$C(\Delta X_t) = \begin{pmatrix} b_1 \Delta X_t, b_1 \geq 0 & \text{iff } \Delta X_t \geq 0 \\ b_2 \Delta X_t, b_2 < 0 & \text{iff} \Delta X_t < 0 \end{pmatrix}$$

and a graphical representation is found in figure (2.1.c). Adjustment costs in this setting are symmetrical only in the case of $b_1 = -b_2$. Again we recognize the signs of asymmetry as costs are relatively large for positive changes. Because even the smallest change in the level of *X* induces positive costs while the marginal costs are constant except at $\Delta X = 0$, it may be optimal for the firm to abstain from adjustments until the associated benefits offset the costs implied.

2.4. d) Lumpy adjustment costs

Because many factor investments yield adjustment costs that must necessarily be partly independent of the actual size of the investments, Hamermesh and Pfann introduce a fourth approximation that incorporates nonconvexity and "lumpiness" in the adjustment cost function. As the authors point out, "The gross, external costs of obtaining plans of acquiring a site and of creating new networks for selling the plant's output all produce some fixed components. Some of the costs of hiring – advertising screening, and training, and others – are up to a point independent of the number of hires.” Further they note that it is certainly possible to “include both lumpy and linear piecewise costs along with a quadratic term to describe adjustment in a more complex manner”. A simple representation of a lumpy adjustment cost function is written as

$$C(\Delta X_t) = k_1 I_1(\Delta X_t) + k_2 I_2(\Delta X_t),$$

and is illustrated in figure (2.1.d). As under piecewise linear costs, the firm will in some situations have incentives to abstain from investing because the associated benefits do not offset the fixed cost the firm will have to face.

2.5. The Effects of Adjustment Costs on Factor Demand

Much of the recent literature on investment behavior and adjustment costs builds on the work of Abel and Eberly (1994), which illustrates how the existence of non-convex adjustment costs can result in optimal investment behavior that is in accordance with what we actually observe. Their paper presents a one-factor model which includes nonconvexity in the adjustment cost function, and derives the firm’s optimal investment policy. The model assumes costs of investments that can be divided into three components; (i) purchase or resale price, (ii) costs of adjustments and (iii) fixed costs. The paper derives an optimal solution of the investment decision that illustrates how nonconvexities under certain conditions can result in the choice of zero adjustment in spite of a change in the optimal level of factor input. The possible existence of a fixed term in the costs of adjustments imposes a more strict condition to be satisfied for investments to occur because an adjustment of any size will trigger a minimum level of costs. Abel and Eberly show that if this fixed term is relatively large, we will expect to see longer periods without changes, and changes to be relatively large once they occur.

As a natural extension of the one-factor model, Eberly and Van Mieghem (1996) introduces a multi-factor analysis where all the included production factors may be subject to adjustment costs, and the degree of fixity is endogenously determined for each factor. The authors use a state space divided into various domains to illustrate possible investment strategies. The domain of main interest is a continuation region where no adjustment of any factor is an optimal solution. As in the case of a one-factor analysis, a continuation region with a non-empty interior stems from investments that are costly to reverse. The shape and size of this region depends on the degree of irreversibility and by possible asymmetries in marginal adjustment costs.

In a similar fashion Dixit (1997) investigates the degree of fixity of two input factors, and introduces a model that generates an endogenous ranking of capital and labor as more or less flexible. Adjustment costs in this setting are assumed to be linear. Many economic models assume fixed levels of capital in the short or semi-long run, while employment may change free of costs. Using a state space illustration with a main interest in the space of no action, Dixit shows that employment dynamics may have many of the same qualities as those which characterize the dynamics of demand for capital, which means that the assumption of free flowing labor and fixed levels of capital in many cases might be unreasonable.

Dixit writes that in practice, decisions regarding employment are not necessarily reversible and hiring and firing cost can be substantial, especially in European economies. The direct costs arising in firing or hiring are examples of adjustment cost related to changes in employment. These include production loss due to interruptions of the production process when reorganizing the work force, search costs and training costs when hiring, workers compensations when firing, and overhead administration costs of both hiring and firing.

One of the most important lessons presented in Dixit's paper is that, no matter which input factor is considered more flexible, "adjustment of the more flexible factor can occur on its own, but that of the less flexible factor occurs less frequently and only in conjunction with a complementary adjustment of the more flexible factor." This should be an important reminder when investigating descriptive statistics of empirical data, and when interpreting results after one-factor analyses.

In a second paper by Abel and Eberly (1998), investment behavior is analyzed in a two-factor model. In accordance with Dixit's conclusions, the authors find that quasi-fixity of one production factor, i.e. capital, may lead to lumpy investments in a second fully flexible input of production. That is, we may see labor hoarding in periods of large investments in capital, even though input of labor does not inhabit any adjustment costs. Again, these results call for caution when interpreting employment dynamics, or other factor dynamics, as investment patterns may be highly dependent on the dynamics of a second quasi-fixed production factor. This should give sufficient motivation for applying a multi-factor analysis when investigating the possible effects of adjustment costs.

Bloom (2009) analyses the impact of uncertainty shocks on output and employment. Because of the possible effect of adjustment costs on aggregate demand for production factors, the model presented jointly estimates both convex and nonconvex costs for labor and capital. Regarding costs of adjustment, Bloom finds that "Ignoring capital adjustment costs is shown to lead to substantial bias while ignoring labor adjustment costs does not".

Contreras (2006) considers possible interrelations in the adjustment cost functions, and the functions are specified to include a component that makes simultaneous investments more or less costly relative to sequential investments. Descriptive statistics indicate "lumpiness" in the investment patterns of both capital and labor, which may or may not be caused by the existence of irreversibility. Also, indications can be seen of asymmetry in capital investment which in turn indicates irreversibility of that factor, and there is a clear pattern of correlation between adjustment spikes. By logit estimation Contreras finds that the probability of inaction in capital adjustment increases with inaction in labor adjustment. The same effect is found in the opposite direction, although the effect is not significant.

3. A Model with Interrelations

3.1. Introduction

When we think that there are characteristics about the adjustment process that make firm level adjustments of both labor and capital lumpy, it would be natural to ask how the demand for two or more production factors actually are affecting each other. When we say interrelations, we simply mean that investment decisions regarding several factors are not made independently. On the contrary, we say that a decision to change the level of one input will affect the investment decisions regarding the remaining factors of production.

As an example, we can consider a simultaneous adjustment of the two production factors, capital and labor. The costs associated with this joint adjustment may increase or decrease the total costs of adjustments relative to a sequential adjustment, where the level of one factor remains unchanged while investment in the other factor is carried out. Whether a simultaneous adjustment will result in a relative cost increase is not really clear. Contreras (2006) finds that an interaction component which makes simultaneous adjustment relatively expensive is best suited to describe the behavior in his sample. Letterie et al. (2004) find that dynamics in employment are significantly correlated with large investments in capital, and vice versa. The authors do not discuss causality, but indicate that interrelations should be considered when estimating the effect of adjustment costs on investment behavior.

3.2. Letterie, Nilsen, Pfann (2009) – "Interrelated Factor Demand with Nonconvex Adjustment Costs."

To show how it is possible to estimate parameters of adjustment cost functions that are both nonconvex and interrelated, we first need to consider a theoretical model with the desired properties. For this purpose, we apply a model introduced by Letterie, Nilsen and Pfann (2009).

The model starts out by considering a firm that employs two factors of production, labor, *L*, and capital, *K* in period *t*, to produce a non-storable good. The firm's objective function is given by

$$V(K_t, L_t, \varepsilon_t) = \max_{I_{t+s}, H_{t+s}} \int_0^\infty E_t \left\{ \pi_t \cdot e^{-\beta s} ds \right\} \qquad [3.1]$$

The discount rate is given by β, with $0 < \beta < 1$. The operating profit of the firm in period *t* is equal to $\pi_t = F(K_t, L_t, \varepsilon_t) - w_t L_t - C(I_t, K_t, H_{t,} L_t)$. The variable w_t denotes the wage paid by the firm to a full time worker. Investment and hiring (or firing) are denoted by I_t and H_t respectively. Sales are given by the expression $F(K_t, L_t, \varepsilon_t)$ where the term ε_t represents a variable capturing randomness in technology or stochastic behavior of the demand conditions the firm is facing. The stochastic term ε_t evolves according to

$$d\varepsilon_t = \mu(\varepsilon_t) + \sigma(\varepsilon_t) \cdot dz \qquad [3.2]$$

where *dz* is a standard Wiener process.

When adjusting the stock of capital or the number of workers the firm incurs adjustment costs defined as:

$$\begin{aligned} C(I_t, K_t, H_t, L_t) = & \left[p_t^{I+} I_t \cdot \mathrm{I}(I_t > 0) + p_t^{I-} I_t \cdot \mathrm{I}(I_t < 0) + \alpha^K + \frac{b^K}{2} \left(\frac{I_t}{K_t} \right)^2 \cdot K_t \right] \cdot \mathrm{I}(I_t \neq 0) \\ & + \left[p_t^{H+} H_t \cdot \mathrm{I}(H_t > 0) + p_t^{H-} H_t \cdot \mathrm{I}(H_t < 0) + \alpha^L + \frac{b^L}{2} \left(\frac{H_t}{L_t} \right)^2 \cdot L_t \right] \cdot \mathrm{I}(H_t \neq 0) \\ & + \alpha^{KL} \cdot \mathrm{I}(I_t \neq 0) \cdot \mathrm{I}(H_t \neq 0) \end{aligned} \qquad [3.3]$$

In the adjustment cost function the indicator function I(.) assumes the value 1 if the condition in brackets is satisfied and equals zero otherwise. We note that the cost of interaction given by the parameter α^{KL} disappears if the firm does not change the level of both factors simultaneously in period *t*. The same term is positive if a joint adjustment would increase the cost relative to a sequential solution. On the other hand, α^{KL} will assume a negative value in cases where a simultaneous strategy will give the firm a relative cost advantage.

We recognize the adjustment costs in this framework as a mix of different structures. The main advantage in the specification above is that the model allows for both asymmetries and nonconvexities. Fixed cost parameters are given by α^{K} and α^{L} and are assumed to be independent of whether the changes of levels of inputs are positive or negative (i.e. symmetric). The model does however, allow for asymmetry when specifying the purchase price of capital as p_t^{I+} , while in the case the firm sells capital the model assumes that the price received for one unit of capital equals p_t^{I-}. Due to irreversibility of investment decisions $p_t^{I+} > p_t^{I-}$, and we note that the price component also includes the actual cost of investment. Linear adjustment costs with respect to hiring and firing are denoted by p_t^{H+} when $H_t > 0$ and p_t^{H-} when $H_t < 0$.

The firm decides the optimal size of the capital stock, K_t, by setting investment I_t, at the appropriate level. Since capital depreciates at rate δ^{K}, the capital stock evolves according to the law of motion

$$dK_t = (I_t - \delta^K K_t) \cdot dt \qquad [3.4]$$

Simultaneously, the firm determines the optimal value for the number of workers L_t, by choosing the desired and hence optimal level of hiring or firing denoted by H_t. The amount of labor evolves according to

$$dL_t = (H_t - \delta^L L_t) \cdot dt \qquad [3.5]$$

where δ^{L} measures the autonomous quit rate of workers.

To obtain the optimal values for I_t and H_t, the objective function is optimized with respect to these decision variables, subject to the laws of motion in equations [3.4] and [3.5].

Before proceeding we note that variables λ_t^K and λ_t^L are conventional marginal values of capital and labor, respectively.

Maximization yields the following first order conditions

$$\lambda_t^K - p_t^{I+} \cdot \mathrm{I}(I_t > 0) - p_t^{I-} \cdot \mathrm{I}(I_t < 0) - b^K \left(\frac{I_t}{K_t} \right) = 0 \quad [3.6]$$

$$\lambda_t^L - p_t^{H+} \cdot \mathrm{I}(H_t > 0) - p_t^{H-} \cdot \mathrm{I}(H_t < 0) - b^L \left(\frac{H_t}{L_t} \right) = 0 \quad [3.7]$$

Hence, optimal amounts of investment and hiring or firing are

$$\frac{I_t}{K_t} = \left(\frac{\lambda_t^K - p_t^I}{b^K} \right) \quad [3.8]$$

$$\frac{H_t}{L_t} = \left(\frac{\lambda_t^L - p_t^H}{b^L} \right) \quad [3.9]$$

Where $p_t^I \equiv p_t^{I+} \cdot \mathrm{I}(I_t > 0) + p_t^{I-} \cdot \mathrm{I}(I_t < 0)$ and $p_t^H \equiv p_t^{H+} \cdot \mathrm{I}(H_t > 0) + p_t^{H-} \cdot \mathrm{I}(H_t < 0)$.

The Effects of Nonconvexities

Due to the presence of fixed costs of adjustment the firm will not always adjust to meet the optimal adjustment levels above. Sometimes it may be optimal to abstain from adjusting capital and or adjusting labor. The threshold equation determining whether to change the stock of capital and or to adjust labor becomes

$$\lambda_t^K I_t + \lambda_t^L H_t \geq C(I_t, K_t, H_t, L_t) \quad [3.10]$$

where the left hand side of the equation measures the expected benefits of changing capital and or labor, whereas the right hand side denotes the cost associated with the firm's decisions. It can be shown that a necessary condition for changing the amount of capital is

$$\left| \lambda_t^L - p_t^I \right| > \sqrt{\frac{2b^K \alpha^K}{K_t}} \equiv A^K \quad [3.11]$$

And a similar condition for hiring or firing is

$$\left|\lambda_t^L - p_t^H\right| > \sqrt{\frac{2b^L\alpha^L}{L_t}} \equiv A^L \qquad [3.12]$$

Equations [3.11] and [3.12] show that the net benefits of adjusting capital and labor do not exceed a certain minimum threshold; the management will decide to abstain from adjusting. These two thresholds are caused by the existence of fixed adjustment costs, but are also affected by the magnitude of convex costs.

Before proceeding with the analysis we consider two possible situations. In the first one, the adjustment costs are structured in a way that makes the costs of joint adjustment large relative to a sequential strategy. This situation would arise if $\alpha^{KL} > 0$. In the opposite case, simultaneous adjustment would reduce the costs of adjustment, and hence increase the possibility of simultaneous investments. This would be the case if $\alpha^{KL} < 0$. Let us now consider the first situation, and assume that both necessary conditions for adjusting capital and labor are satisfied as given in equations [3.11] and [3.12]. It can be shown that it is worth also adjusting the stock of capital (given that adjusting labor yields a higher value of the firm if only one input needs to be selected) as soon as

$$\frac{1}{2b^K}\left(\lambda_t^K - p_t^I\right)^2 K_t \geq \alpha^K + \alpha^{KL} \qquad [3.13]$$

Similarly, labor will also be adjusted (given that changing capital yields a higher firm value if only one input is selected) as soon as

$$\frac{1}{2b^L}\left(\lambda_t^L - p_t^H\right)^2 L_t \geq \alpha^L + \alpha^{KL} \qquad [3.14]$$

Hence, the boundaries determining when the firm will adjust both factors of production are

$$\left|\lambda_t^K - p_t^I\right| > \sqrt{\frac{2b^K\left(\alpha^K + \alpha^{KL}\right)}{K_t}} \equiv B^K \qquad [3.15]$$

$$\left|\lambda_t^L - p_t^H\right| > \sqrt{\frac{2b^L\left(\alpha^L + \alpha^{KL}\right)}{L_t}} \equiv B^L \qquad [3.16]$$

In the case when $\alpha^{KL} < 0$, the firm will actually incur a cost reduction by making joint adjustments compared to a sequential strategy. In this situation, the threshold values B^K and B^L will be lower than the values of A^K and A^L respectively. More specifically we can say that if $\alpha^L + \alpha^{KL} \leq 0$, this would mean that it will always be optimal for the firm to change the level of employment in every period it changes the level of capital. This is true because the effect of the fixed term α^L is completely balanced by the cost advantage represented by $\alpha^{KL} < 0$. Similarly it will always be optimal to change the level of capital input together with employment changes if $\alpha^K + \alpha^{KL} \leq 0$. If we consider a situation where both conditions $\alpha^L + \alpha^{KL} \leq 0$ and $\alpha^K + \alpha^{KL} \leq 0$ are satisfied, a joint investment strategy will always be preferred over a sequential strategy.

4. Model Set-up and Parameterization

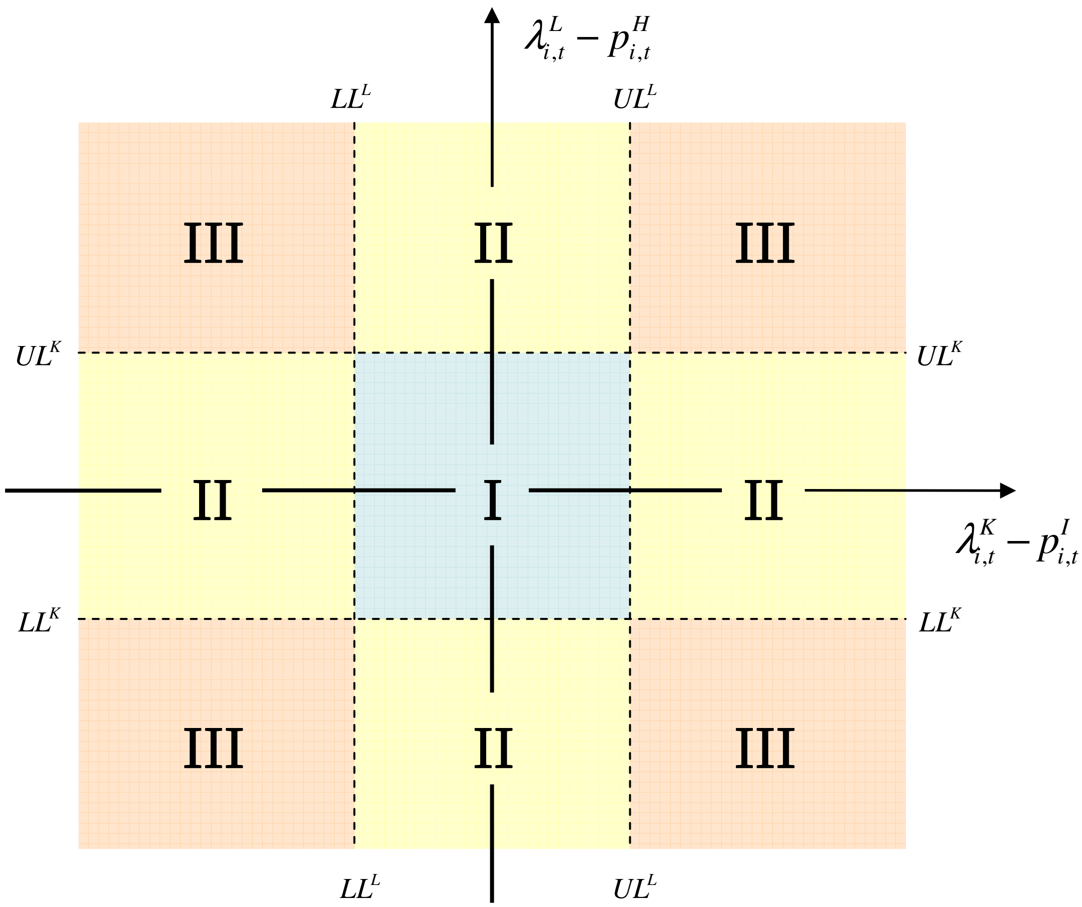
4.1. Introduction

To empirically estimate the parameters of the theoretical model presented in section 3.2, we need to develop a suitable estimation technique. We wish to investigate how firms in a sample make their investment decisions in period *t*. To simplify the problem we say that every firm decides between three options per input factor. These options are positive investment, negative investment and no investment. Decisions regarding two input factors in period *t* will place the strategy employed by the firm in one of three investment regimes: No changes (I), changes in one factor only (II), and changes in both factors (III). Figure (4.1) illustrates how the choices concerning adjustment of capital and labor can be divided into several action spaces or investment regimes. Regime I represents the case of no investment, regime II is the case of change in only one of the factors, while regime III represents the case of simultaneous change. The X- and Y-axis measure the marginal benefit of labor and capital respectively. With large changes in the shadow values $\lambda_{i,t}^{j}$ or purchase prize $p_{i,t}^{j}$, the optimal strategy of the firm will move away from no investment in regime I and over to one of the other regimes. The optimal investment decisions for each factor change as the values of $\lambda_{i,t}^{L} - p_{i,t}^{H}$ and $\lambda_{i,t}^{K} - p_{i,t}^{I}$ move across certain threshold values. These are denoted $LL_{i,t}^{L}, UL_{i,t}^{L}, LL_{i,t}^{K}$ and $UL_{i,t}^{K}$ in figure 4.1. For simplicity, the figure illustrates a situation without an interrelation cost parameter, and thresholds correspond directly to the sequential factor adjustment thresholds from equations *equation reference goes here* That means

$$LL_{i,t}^{L} = -A_{i,t}^{L} = -\sqrt{\frac{2b^{L}\alpha^{L}}{L_{i,t}}}, \qquad UL_{i,t}^{L} = A_{i,t}^{L} = \sqrt{\frac{2b^{L}\alpha^{L}}{L_{i,t}}}$$

$$LL_{i,t}^{K} = -A_{i,t}^{K} = -\sqrt{\frac{2b^{K}\alpha^{K}}{K_{i,t}}}, \qquad UL_{i,t}^{K} = A_{i,t}^{K} = \sqrt{\frac{2b^{K}\alpha^{K}}{K_{i,t}}}$$

FIGURE 4.1: INVESTMENT REGIMES



The probabilities of observing the different choices of investment regimes will depend on threshold levels derived in equations [3.11] and [3.12], which we have found to be dependent on characteristics in the adjustment cost functions. Since we are considering choices concerning two production factors simultaneously, we can derive a joint limited dependent variable model much similar to a seemingly unrelated bivariate ordered probit model. The following sections will show how the model can be derived, and how it can be estimated by a Maximum Likelihood routine.

4.2. Maximum Likelihood [2]

The starting point of maximum likelihood estimation is the assumption that the (conditional) distribution of an observed phenomenon (the endogenous variable) is known, except for a finite number of unknown parameters. These parameters will be estimated by taking those values for them that give the observed values with the highest probability, the highest likelihood.

To enable maximum likelihood estimation we need to make an assumption about the shape of the distribution of the error terms. The most common assumption is that ε_i is normally and independently distributed (n.i.d.) with mean zero and variance σ^2, or $\varepsilon_i \sim NID(0,\sigma^2)$.

Let the density or probability mass function be given by $f\left(y_i \middle| x_i,\theta\right)$, where θ is a K-dimensional vector of unknown parameters, and assume that observations are mutually independent. In this situation, the joint probability mass function of the sample $y_1,..., y_N$ is given by

$$f\left(y_1,..., y_N \left| X;\theta\right.\right) = \prod_{i=1}^{N} f\left(y_i \left| x_i;\theta\right.\right) \qquad [4.1]$$

The likelihood function for the entire sample is then given by

$$L\left(\theta \middle| y, X\right) = \prod_{i=1}^{N} L_i\left(\theta \middle| y_i, x_i\right) = \prod_{i=1}^{N} f\left(y_i \left| x_i;\theta\right.\right) \qquad [4.2]$$

The ML-estimator $\hat{\theta}$ for θ is the solution to

$$\max_{\theta} \log L\left(\theta\right) = \max_{\theta} \sum_{i=1}^{N} \log L_i(\theta) \qquad [4.3]$$

, where $\log L\left(\theta\right)$ is the loglikelihood function.

Provided that the likelihood function is correctly specified, it can be shown under weak regularity conditions that:

[2] This subsection is taken directly from Verbeek (2008), where a more detailed discussion can be found.

1. The maximum likelihood estimator is consistent for θ (plim $\hat{\theta} = \theta$)
2. The maximum likelihood estimator is asymptotically efficient
3. The maximum likelihood estimator is asymptotically normally distributed, according to $\sqrt{N}\left(\hat{\theta} - \theta\right) \rightarrow Norm(0, V)$.

4.3. Deriving the Likelihood Function

We have seen how a theoretical model that specifies an adjustment cost function for two input factors is capable of predicting investment behavior with lumps and bumps. This analysis will consider the investment decisions for labor and capital in the theoretical model as two discrete variables which we can categorize as limited dependent variables. Now we need to develop a suitable likelihood function to be maximized by Maximum Likelihood.

As the individual firm sees marginal values change, optimal adjustment strategies might also be altered. The econometrician will be able to observe firms' investment decisions. However, marginal values $\lambda_{i,t}^{L}$ and $\lambda_{i,t}^{K}$ are not observable. As in any standard limited dependent variable model we say that the observed variable takes its values conditional on an unobserved latent variable. In other words, factor adjustments depend on factor profitability, which we can not observe.

We start out by considering functions for the latent marginal factor values

$$\lambda_{i,t}^{L} = \beta_0^{L} + \beta_1^{L} Z_{i,t}^{L} + p_t^{H} + \varepsilon_{i,t}^{L} \qquad [4.4]$$

$$\lambda_{i,t}^{K} = \beta_0^{K} + \beta_1^{K} Z_{i,t}^{K} + p_t^{I} + \varepsilon_{i,t}^{K} \qquad [4.5]$$

Equations [4.4] and [4.5] tell us that the shadow values $\lambda_{i,t}^{L}$ and $\lambda_{i,t}^{K}$ are dependent variables and functions of one explanatory variable $Z_{i,t}^{j}$, in addition to a constant term, $\left(\beta_0^j + p_t^j\right)$ and stochastic error terms $\varepsilon_{i,t}^{L}$ and $\varepsilon_{i,t}^{K}$ that include all variables affecting marginal values that are not observable. We note that we will not be able to identify β_0^j and p_t^j as these two parameters will be components of an estimated constant term. Also, we should be careful about defining explanatory variables $Z_{i,t}^{L}$ and $Z_{i,t}^{K}$. A standard approach in the factor adjustment literature is that these are variables of output-to-labor and output-to-capital ratios respectively. This is based on a derivation shown in appendix A. The approach traditionally used is unfortunately based on assumptions that are somewhat in contrast what motivates this thesis, and the problem is discussed in more detail in the appendix. For now we continue with our derivation of the likelihood function.

To derive the likelihood function for use in an ML estimation, we can start out by considering figure (4.1). As we observe firms' investment decisions we define our limited dependent variable, $Y_{i,t}$, which describes firm i's choice of investment strategy I, II or III in period t.

$$
\begin{aligned}
Y_{i,t} &= \text{Regime I} \quad && \textit{if} \quad \text{no adjustment} \\
Y_{i,t} &= \text{Regime II} \quad && \textit{if} \quad \text{adjustment of one factor} \\
Y_{i,t} &= \text{Regime III} \quad && \textit{if} \quad \text{adjustment of both factors}
\end{aligned}
$$

Figure 4.1 shows how the investment regimes are limited by upper and lower thresholds for which we found expressions in chapter 3. We also remember that these threshold values apply to the latent variables $\lambda_{i,t}^{L}$ and $\lambda_{i,t}^{K}$ net of factor prices p_t^H and p_t^I. If the net marginal value does not exceed one of the respective thresholds, we will see no adjustment of that factor. Using equations [4.4] and [4.5], we can express the investment variable $Y_{i,t}$ conditional on the determinants of latent marginal values.

$$Y_{i,t} = \text{Regime I} \quad \textbf{if} \quad LL_{i,t}^L < \left(\beta_0^L + p_t^H\right) + \beta_1^L Z_{i,t}^L + \varepsilon_{i,t}^L < UL_{i,t}^L$$
$$\textbf{and } LL_{i,t}^K < \left(\beta_0^K + p_t^I\right) + \beta_1^K Z_{i,t}^K + \varepsilon_{i,t}^K < UL_{i,t}^K$$

$$Y_{i,t} = \text{Regime II} \quad \textbf{if} \quad LL_{i,t}^L < \left(\beta_0^L + p_t^H\right) + \beta_1^L Z_{i,t}^L + \varepsilon_{i,t}^L < UL_{i,t}^L$$
$$\textbf{and } \left(\beta_0^K + p_t^I\right) + \beta_1^K Z_{i,t}^K + \varepsilon_{i,t}^K > UL_{i,t}^K \quad \textbf{or} \quad \left(\beta_0^K + p_t^I\right) + \beta_1^K Z_{i,t}^K + \varepsilon_{i,t}^K < LL_{i,t}^K$$

Or

$$Y_{i,t} = \text{Regime II} \quad \textbf{if} \quad LL_{i,t}^K < \left(\beta_0^K + p_t^I\right) + \beta_1^L Z_{i,t}^K + \varepsilon_{i,t}^K < UL_{i,t}^K$$
$$\textbf{and } \left(\beta_0^L + p_t^H\right) + \beta_1^L Z_{i,t}^L + \varepsilon_{i,t}^L > UL_{i,t}^L \quad \textbf{or} \quad \left(\beta_0^L + p_t^H\right) + \beta_1^L Z_{i,t}^L + \varepsilon_{i,t}^L < LL_{i,t}^L$$

$$Y_{i,t} = \text{Regime III} \quad \textbf{if} \quad \left(\beta_0^L + p_t^H\right) + \beta_1^L Z_{i,t}^L + \varepsilon_{i,t}^L > UL_{i,t}^L \quad \textbf{or} \quad \left(\beta_0^L + p_t^H\right) + \beta_1^L Z_{i,t}^L + \varepsilon_{i,t}^L < LL_{i,t}^L$$
$$\textbf{and } \left(\beta_0^K + p_t^I\right) + \beta_1^K Z_{i,t}^K + \varepsilon_{i,t}^K > UL_{i,t}^K \quad \textbf{or} \quad \left(\beta_0^K + p_t^I\right) + \beta_1^K Z_{i,t}^K + \varepsilon_{i,t}^K < LL_{i,t}^K$$

Given our simplification by assuming the interrelation cost parameter $\alpha^{KL} = 0$ thresholds are given by.

$$LL_{i,t}^L = -A_{i,t}^L = -B_{i,t}^L = -\sqrt{\frac{2b^L\alpha^L}{L_{i,t}}}, \qquad UL_{i,t}^L = A_{i,t}^L = B_{i,t}^L = \sqrt{\frac{2b^L\alpha^L}{L_{i,t}}}$$
$$LL_{i,t}^K = -A_{i,t}^K = -B_{i,t}^K = -\sqrt{\frac{2b^K\alpha^K}{K_{i,t}}}, \qquad UL_{i,t}^K = A_{i,t}^K = B_{i,t}^K = \sqrt{\frac{2b^K\alpha^K}{K_{i,t}}}$$

[4.6]

The likelihood function to be used in the ML estimation follows a standard setup for discrete variables. We have the probability for one observation as

$$f = \Pr\left(Y_{i,t} \middle| \mathbf{x}_{i,t}, \boldsymbol{\theta}\right) \qquad Y_i \in (1,2,3)$$

Where $x_{i,t}$ denotes all explanatory variables for firm *i* at time *t*, and $\boldsymbol{\theta}$ is a vector of all estimated parameters.

Accordingly we can write a likelihood function for the entire sample as

$$L = \prod_{i=1}^{N} f(.) = \prod_{i=1}^{N} \Pr\left(Y_{i,t} \middle| \mathbf{x}_{i,t}, \boldsymbol{\theta}\right) \qquad [4.7]$$

, and the logarithmic likelihood function can then be written as

$$\text{Log } L = \sum_{i=1}^{N} \text{Log } \Pr\left(Y_{i,t} \middle| \mathbf{x}_{i,t}, \boldsymbol{\theta}\right) \qquad [4.8]$$

$$\begin{aligned}\text{Log } L = &\sum_{t=1}^{T} \sum_{i \in \Omega^{I}} \log \Pr\left(Y_{i,t} = \text{Regime I} \middle| \mathbf{x}_{i,t}, \boldsymbol{\theta}\right) + \sum_{t=1}^{T} \sum_{i \in \Omega^{II}} \log \Pr\left(Y_{i,t} = \text{Regime II} \middle| \mathbf{x}_{i,t}, \boldsymbol{\theta}\right) \\ &+ \sum_{t=1}^{T} \sum_{i \in \Omega^{III}} \log \Pr\left(Y_{i,t} = \text{Regime III} \middle| \mathbf{x}_{i,t}, \boldsymbol{\theta}\right)\end{aligned} \qquad [4.9]$$

The logarithmic likelihood function tells us that the probability of observing firm i in its respective investment regime at time *t* is conditional on all explanatory variables $\mathbf{x}_{i,t}$ and estimated parameters $\boldsymbol{\theta}$. Since the observed adjustment variable $Y_{i,t}$ is a function of two latent variables, it is also affected by the explanatory variables $\mathbf{x}_{i,t}$ and parameters $\boldsymbol{\theta}$ from both equations [4.4] and [4.5] simultaneously. Given $\mathbf{x}_{i,t}$ and estimation of $\boldsymbol{\theta}$, the residuals $\varepsilon_{i,t}^{L}$ and $\varepsilon_{i,t}^{K}$ will determine the probability of each observed $Y_{i,t}$. To be able to consider interrelations in the two latent equations, we need to specify a functional form for the two processes which allows for correlation. We assume the following about our error terms. [3]

[3] The presentation of the bivariate normal distribution function follows Greene (2008).

$$E\left[\varepsilon_{i,t}^{L} \middle| Z_{i,t}^{L}, Z_{i,t}^{L}\right] = E\left[\varepsilon_{i,t}^{K} \middle| Z_{i,t}^{K}, Z_{i,t}^{K}\right] = 0 \quad [4.10]$$

$$Var\left[\varepsilon_{i,t}^{L} \middle| Z_{i,t}^{L}, Z_{i,t}^{L}\right] = Var\left[\varepsilon_{i,t}^{K} \middle| Z_{i,t}^{K}, Z_{i,t}^{K}\right] = 1 \quad [4.11]$$

$$Cov\left[\varepsilon_{i,t}^{L}, \varepsilon_{i,t}^{K} \middle| Z_{i,t}^{L}, Z_{i,t}^{K}\right] = \rho \quad [4.12]$$

In addition, we assume that $\varepsilon_{i,t}^{L}$ and $\varepsilon_{i,t}^{K}$ are normally distributed.

Now, to estimate probability contributions for each observation we need to express the probabilities of estimated error terms. For this purpose we the use the bivariate normal cumulative distribution function

$$\Pr(X_1 < x_1, X_2 < x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \phi_2\left(z_1, z_2, \rho\right) dz_1 dz_2 \quad [4.13]$$

Which is defined as

$$\Phi_2(x_1, x_2, p).$$

The simultaneous distribution function is

$$\phi_2\left(x_1, x_2, \rho\right) = \frac{\exp\left(-(1/2)\left(x_1^2 + x_2^2 - 2\rho x_1 x_2\right) \big/ \left(1-\rho^2\right)\right)}{2\pi\left(1-\rho^2\right)^{1/2}} \quad [4.14]$$

This distribution provides for us a way to identify the volumes of the three investment regimes in figure (4.1). The upper limits that are named x_1 and x_2 equation [4.13] will be given values according to our underlying theoretical model. Using the bivariate normal distribution function, it can be shown that the probabilities or likelihood contribution for each individual observation can be written as:

$$\begin{aligned}\Pr\left(Y_{i,t}=\text{Regime I}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right)=&\,\Phi_2\left[ul^L_{i,t},ul^K_{i,t},\rho\right]-\Phi_2\left[ll^L_{i,t},ul^K_{i,t},\rho\right]\\&-\Phi_2\left[ul^L_{i,t},ll^K_{i,t},\rho\right]+\Phi_2\left[ll^L_{i,t},ll^K_{i,t},\rho\right]\end{aligned}\qquad[4.15]$$

$$\begin{aligned}\Pr\left(Y_{i,t}=\text{Regime II}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right)=&\left(\Phi_2\left[ll^L_{i,t},ul^K_{i,t},\rho\right]-\Phi_2\left[ll^L_{i,t},ll^K_{i,t},\rho\right]\right)\\&+\left(\Phi\left[ul^L_{i,t}\right]-\Phi\left[ll^L_{i,t}\right]-\Phi_2\left[ul^L_{i,t},ul^K_{i,t},\rho\right]+\Phi_2\left[ul^L_{i,t},ul^K_{i,t},\rho\right]\right)\\&+\left(\Phi\left[ul^K_{i,t}\right]-\Phi_2\left[ul^L_{i,t},ul^K_{i,t},\rho\right]-\Phi\left[ll^K_{i,t}\right]+\Phi_2\left[ul^L_{i,t},ll^K_{i,t},\rho\right]\right)\\&+\left(\Phi_2\left[ul^L_{i,t},ul^K_{i,t},\rho\right]-\Phi_2\left[ll^L_{i,t},ll^K_{i,t},\rho\right]\right)\end{aligned}$$

[4.16]

$$\begin{aligned}\Pr\left(Y_{i,t}=\text{Regime III}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right)=&\left(\Phi\left[ll^L_{i,t}\right]-\Phi_2\left[ll^L_{i,t},ll^K_{i,t},\rho\right]\right)\\&+\left(1-\Phi\left[ul^L_{i,t}\right]-\Phi\left[ul^K_{i,t}\right]+\Phi_2\left[ul^L_{i,t},ul^K_{i,t},\rho\right]\right)\\&+\Phi_2\left[ll^L_{i,t},ll^K_{i,t},\rho\right]\\&+\left(\Phi\left[ll^K_{i,t}\right]-\Phi_2\left[ul^L_{i,t},ll^K_{i,t},\rho\right]\right)\end{aligned}\qquad[4.17]$$

where $\quad ll^L_{i,t}=LL^L_{i,t}-\beta^L_0-\beta^L_1Z^L_{i,t}-p^H_{i,t},\qquad ul^L_{i,t}=UL^L_{i,t}-\beta^L_0-\beta^L_1Z^L_{i,t}-p^H_{i,t}$

$$ll^K_{i,t}=LL^K_{i,t}-\beta^K_0-\beta^K_1Z^K_{i,t}-p^I_{i,t},\qquad ul^K_{i,t}=UL^K_{i,t}-\beta^K_0-\beta^K_1Z^K_{i,t}-p^I_{i,t}$$

[4.18]

Inserting equations [4.15] - [4.18] into equation [4.9] completes the derivation of a likelihood function for the case of no interrelation cost, $\alpha^{KL}=0$.

4.4. Adjusting For the Existence of $\alpha^{KL} \neq 0$

Up to this point we have ignored the contents of our thresholds in figure (4.1). Since we have found threshold levels for both sequential and simultaneous adjustment, we remember that these are not identical, except for the special case where $\alpha^{KL} = 0$ (no interrelation cost parameter). On the contrary, threshold levels may be closer to or further away from zero depending on the sign of α^{KL}, which makes our expressions for the action spaces ambiguous.

The figures (4.2) and (4.3) illustrate possible effects of the interaction term α^{KL}. Figure (4.2) shows the grid of investment regimes in the case of a positive interaction term, making joint adjustments relatively costly. In the second figure, the sign of the interaction term is negative, which causes an opposite effect. Footers *A* and *B* denote thresholds relevant for separate or joint adjustments respectively.

Now we have two sets of lower and upper level thresholds for each of the two production factors. One for sequential, and one for simultaneous adjustments. As can clearly be seen from the two diagrams, a negative α^{KL} makes joint adjustments more likely to occur as the size of regime III clearly grows in figure (4.3). Given what we found in section 3, this is perhaps old news. However, this effect has important implications for our likelihood function as we need to specify probability expressions conditional on the sign of α^{KL}.

FIGURE 4.2: INVESTMENT REGIMES, $\alpha^{KL} > 0$

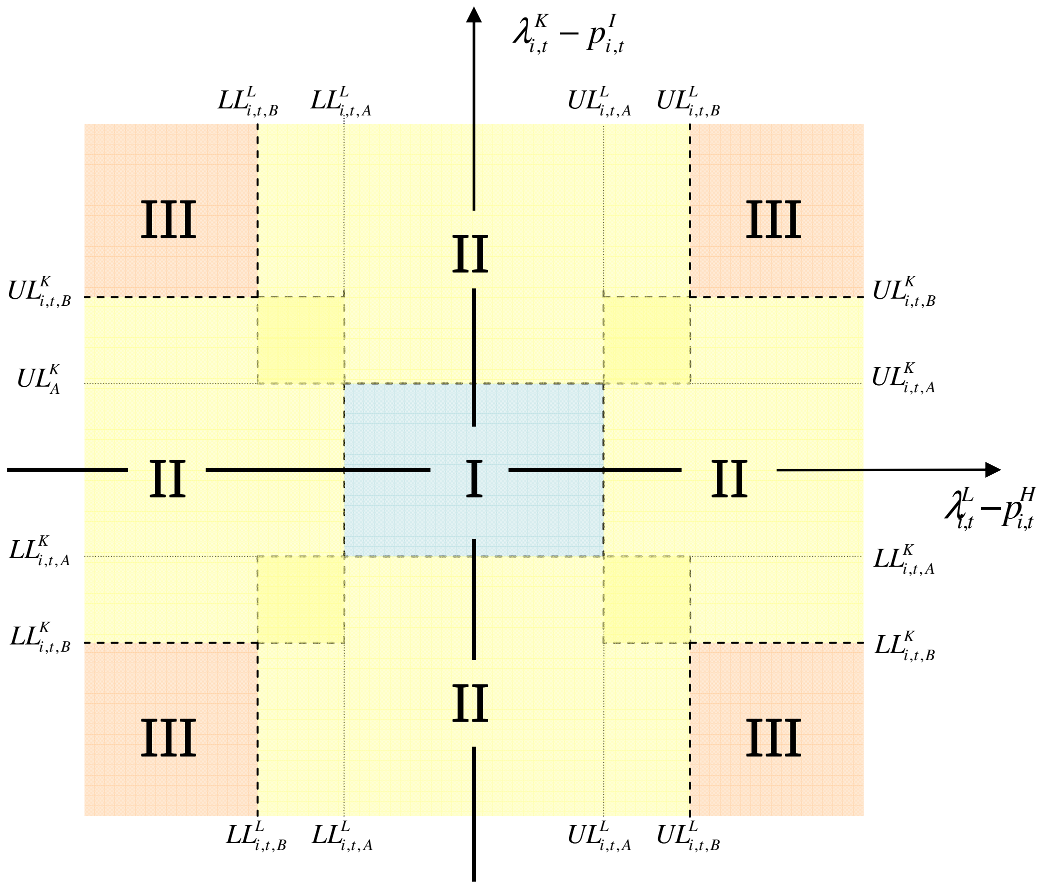
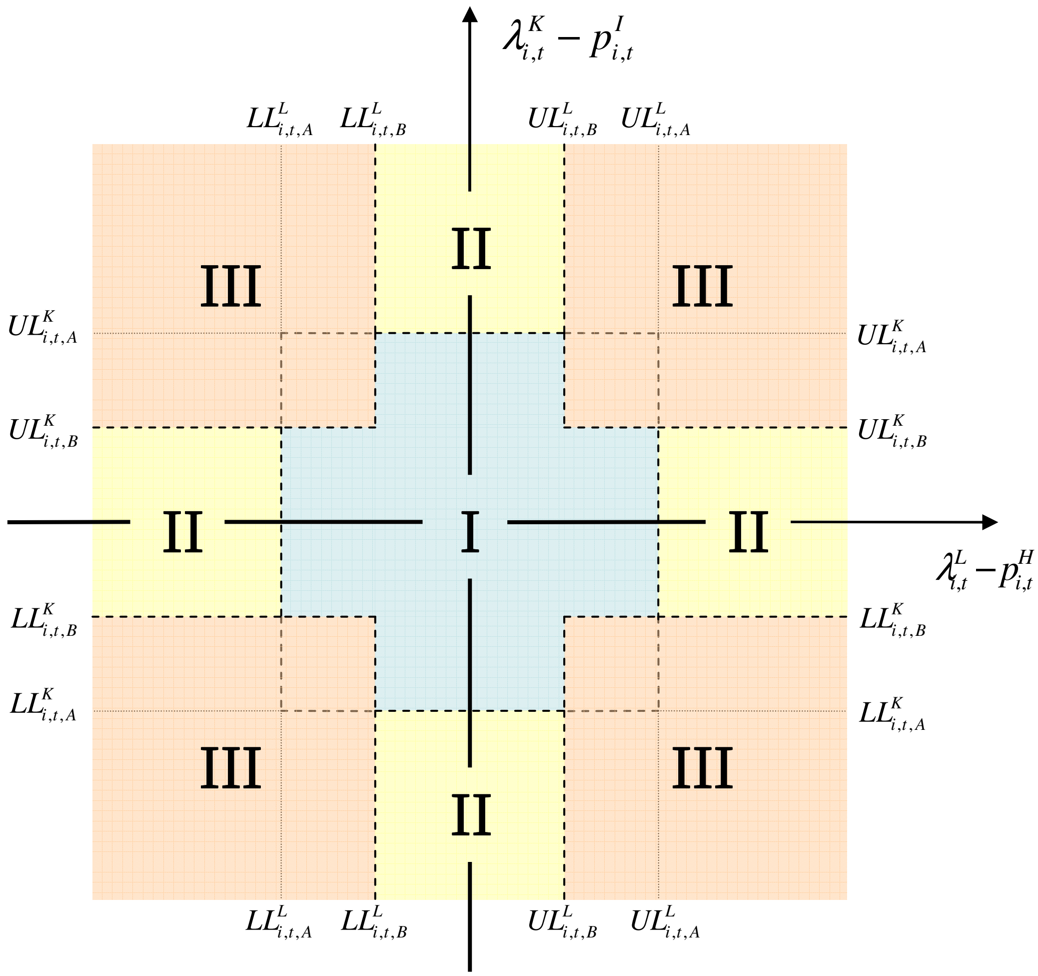


FIGURE 4.3: INVESTMENT REGIMES, $\alpha^{KL} < 0$



First, we notice how the four spaces that make up regime III are limited by the same thresholds in both cases, even though they have different sizes. This is unfortunately not the case for the volume of the inaction space, regime I, which in figure (4.2) is limited only by LL^L_A, UL^L_A, LL^K_A and UL^K_A, but in figure (4.3) also is affected by LL^L_B, UL^L_B, LL^K_B and UL^K_B. The same effect also applies to the action spaces of regime II, which clearly overlap in figure (4.2) but not in figure (4.3). Our likelihood function, therefore, needs to express the volumes of regimes I and II conditional on the sign of α^{KL}. We start out with the same setup as before.

$$\text{Log } L = \sum_{t=1}^{T}\sum_{i\in\Omega^{I}} \log \ \Pr\left(Y_{i,t} = \text{Regime I}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right) + \sum_{t=1}^{T}\sum_{i\in\Omega^{II}} \log \ \Pr\left(Y_{i,t} = \text{Regime II}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right)$$
$$+\sum_{t=1}^{T}\sum_{i\in\Omega^{III}} \log \ \Pr\left(Y_{i,t} = \text{Regime III}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right)$$

[4.9]

This time, however, we want to consider the effect of α^{KL}, and therefore we need new to revise the probability expressions. Using the fact that the volume of all the action spaces summarize to unity, we can write new unambiguous action spaces as the following conditional equations

$$\Pr\left(Y_{i,t} = \text{Regime I}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right) = \begin{cases} \Pr\left(Y_{i,t} = \text{Regime I}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right)^{*} & \text{if } \ \alpha^{KL} > 0 \\ 1 - \begin{bmatrix} \Pr\left(Y_{i,t} = \text{Regime III}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right) \\ +\Pr\left(Y_{i,t} = \text{Regime II}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right) \end{bmatrix} & \text{if } \ \alpha^{KL} < 0 \end{cases} \quad [4.19]$$

$$\Pr\left(Y_{i,t} = \text{Regime II}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right) = \begin{cases} \Pr\left(Y_{i,t} = \text{Regime II}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right)^{*} & \text{if } \ \alpha^{KL} < 0 \\ 1 - \begin{bmatrix} \Pr\left(Y_{i,t} = \text{Regime III}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right) \\ +\Pr\left(Y_{i,t} = \text{Regime I}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right) \end{bmatrix} & \text{if } \ \alpha^{KL} < 0 \end{cases} \quad [4.20]$$

$$\Pr\left(Y_{i,t} = \text{Regime III}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right) = \Pr\left(Y_{i,t} = \text{Regime III}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right)^{*} \quad [4.21]$$

In accordance with what we observe from the figures, the sign of α^{KL} does not alter the boundaries of action spaces with simultaneous adjustments (Regime III). The possibility of $\alpha^{KL} \neq 0$ does however, make it necessary to separate between two sets of upper and lower level thresholds. One that applies when $\alpha^{KL} < 0$, and one that applies when $\alpha^{KL} > 0$. In the firs case the probability expression for Regime II is given as a residual, while in the latter, Regime I is a residual. The full expressions for probabilities marked with asterisks are given below.

$$\begin{aligned}\Pr\left(Y_{i,t}=\text{Regime I}\middle|x_{i,t},\theta\right)^* &= \Phi_2\left[ul_{i,t,A}^L,ul_{i,t,A}^K,\rho\right]-\Phi_2\left[ll_{i,t,A}^L,ul_{i,t,A}^K,\rho\right] \\ &\quad -\Phi_2\left[ul_{i,t,A}^L,ll_{i,t,A}^K,\rho\right]+\Phi_2\left[ll_{i,t,A}^L,ll_{i,t,A}^K,\rho\right]\end{aligned} \qquad [4.22]$$

$$\begin{aligned}\Pr\left(Y_{i,t}=\text{Regime II}\middle|x_{i,t},\theta\right)^* &= \left(\Phi_2\left[ll_{i,t,A}^L,ul_{i,t,B}^K,\rho\right]-\Phi_2\left[ll_{i,t,A}^L,ll_{i,t,B}^K,\rho\right]\right) \\ &\quad +\left(\Phi\left[ul_{i,t,B}^L\right]-\Phi\left[ll_{i,t,B}^L\right]-\Phi_2\left[ul_{i,t,B}^L,ul_{i,t,A}^K,\rho\right]+\Phi_2\left[ll_{i,t,B}^L,ul_{i,t,A}^K,\rho\right]\right) \\ &\quad +\left(\Phi\left[ul_{i,t,B}^K\right]-\Phi_2\left[ul_{i,t,A}^L,ul_{i,t,B}^K,\rho\right]-\Phi\left[ll_{i,t,B}^K\right]+\Phi_2\left[ul_{i,t,A}^L,ll_{i,t,B}^K,\rho\right]\right) \\ &\quad +\left(\Phi_2\left[ul_{i,t,B}^L,ul_{i,t,A}^K,\rho\right]-\Phi_2\left[ll_{i,t,B}^L,ll_{i,t,A}^K,\rho\right]\right)\end{aligned}$$

[4.23]

$$\begin{aligned}\Pr\left(Y_{i,t}=\text{Regime III}\middle|\mathbf{x}_{i,t},\boldsymbol{\theta}\right)^* &= \left(\Phi\left[ll_{i,t,B}^L\right]-\Phi_2\left[ll_{i,t,B}^L,ul_{i,t,B}^K,\rho\right]\right) \\ &\quad +\left(1-\Phi\left[ul_{i,t,B}^L\right]-\Phi\left[ul_{i,t,B}^K\right]+\Phi_2\left[ul_{i,t,B}^L,ul_{i,t,B}^K,\rho\right]\right) \\ &\quad +\Phi_2\left[ll_{i,t,B}^L,ll_{i,B}^K,\rho\right]+\left(\Phi\left[ll_{i,t,B}^K\right]-\Phi_2\left[ul_{i,t,B}^L,ll_{i,t,B}^K,\rho\right]\right)\end{aligned}$$

[4.24]

Where

$$ll_{i,t,A}^L=LL_{i,t,A}^L-\beta_0^L-\beta_1^L Z_{i,t}^L-p_{i,t}^H, \qquad ul_{i,t,A}^L=UL_{i,t,A}^L-\beta_0^L-\beta_1^L Z_{i,t}^L-p_{i,t}^H$$

$$ll_{i,t,A}^K=LL_{i,t,A}^K-\beta_0^K-\beta_1^K Z_{i,t}^K-p_{i,t}^I, \qquad ul_{i,t,A}^K=UL_{i,t,A}^K-\beta_0^K-\beta_1^K Z_{i,t}^K-p_{i,t}^I$$

$$ll_{i,t,B}^L=LL_{i,t,B}^L-\beta_0^L-\beta_1^L Z_{i,t}^L-p_{i,t}^H, \qquad ul_{i,t,B}^L=UL_{i,t,B}^L-\beta_0^L-\beta_1^L Z_{i,t}^L-p_{i,t}^H$$

$$ll_{i,t,B}^K=LL_{i,t,B}^K-\beta_0^K-\beta_1^K Z_{i,t}^K-p_{i,t}^I, \qquad ul_{i,t,B}^K=UL_{i,t,B}^K-\beta_0^K-\beta_1^K Z_{i,t}^K-p_{i,t}^I$$

[4.25]

, and

$$LL_{i,t,A}^L=-A_{i,t}^L=-\sqrt{\frac{2b^L\alpha^L}{L_{i,t}}}, \qquad UL_{i,t,A}^L=A_{i,t}^L=\sqrt{\frac{2b^L\alpha^L}{L_{i,t}}}$$

$$LL_{i,t,A}^K=-A_{i,t}^K=-\sqrt{\frac{2b^K\alpha^K}{K_{i,t}}}, \qquad UL_{i,t,A}^K=A_{i,t}^K=\sqrt{\frac{2b^K\alpha^K}{K_{i,t}}}$$

$$LL_{i,t,B}^L=-B_{i,t}^L=-\sqrt{\frac{2b^L\alpha^L}{L_{i,t}}}, \qquad UL_{i,t,B}^L=B_{i,t}^L=\sqrt{\frac{2b^L\alpha^L}{L_{i,t}}}$$

$$LL_{i,t,B}^K=-B_{i,t}^K=-\sqrt{\frac{2b^K\alpha^K}{K_{i,t}}}, \qquad UL_{i,t,B}^K=B_{i,t}^K=\sqrt{\frac{2b^K\alpha^K}{K_{i,t}}} \qquad [4.26]$$

By inserting equations [4.19] - [4.26] into equation [4.9] we obtain a logarithmic likelihood function for a whole sample which now contains probability expressions that are valid in both figure (4.2) and (4.3). In other words, it should enable us to estimate a model with both non-convexities and interrelations in the adjustment cost functions of two factors, regardless of the sign of the interrelation parameter α^{KL}. We must now consider how this likelihood function can help us identify the different underlying adjustment cost parameters.

4.5. Parameterization

In section 4.3, a loglikelihood function is derived for use in a Maximum Likelihood estimation routine. For the sake of exact identification, we have also seen the need of altering the assumed adjustment cost function. Before I go on to present some results, I will try to give a brief summary of how estimates of the cost parameters can be found by manually defining the structure of the likelihood function, and maximizing the equations using STATA's ML command. First, I will present the basic idea behind the programming structure, after which I will give a little more detailed explanation on which estimates can be found through such a procedure.

4.5.1. The basic set-up

To better understand the issues of estimating the likelihood function, I start out by considering a basic set-up behind programming the Maximum Likelihood function.

1st step: Define variables to represent underlying cost parameters, and define equations for the respective threshold values. This is where we include our adjustment cost parameters and show how they affect the respective thresholds.

2nd step: Define probability expressions for the three investment regimes, conditional on the sign of α^{KL}, using threshold values defined in step 1. This is done by expressing each probability as the sum of integrals of the bivariate normal distribution. The probabilities were derived as unambiguous expressions in section 4.4, where the sign of α^{KL} was identified as a key issue for specifying the likelihood function consistent with the theoretical model, and figures (4.2) and (4.3)

3rd step: Define the final logarithmic likelihood function given by the sum of probability expressions in step 3.

4th step: Maximize the function given in step 4, by setting cost parameters and coefficients to values that jointly maximize the probability of the investment behavior in a simulated dataset. The intuition behind the procedure is to maximize the joint probability of the observations through the setting of thresholds in addition to the coefficients of the lambdas. These thresholds are defined by cost parameters, which allows for maximizing the function with respect to the parameters from the theoretical model.

4.5.2. Setting Cost Parameters

Although getting estimates for all components of the adjustment cost function now seems to be within reach, the observant reader will notice that it may not be possible to identify the size of all parameters, and indeed this is the case for this procedure. When we derived threshold values in section 3.1, we found that the marginal value of an additional unit of one production factor net of the purchase price must be larger than the threshold value *A* or *B*. The appropriate threshold (A or B) depended on the choice of sequencing. Since upper and lower thresholds are identical, as seen from equations given below, we only define four different threshold variables. Thus, we write

$$LL^{L}_{i,t,A} = -A^{L}_{i,t}\text{, } UL^{L}_{i,t,A} = A^{L}_{i,t}\text{, } LL^{L}_{i,t,B} = -B^{L}_{i,t} \text{ and } UL^{L}_{i,t,B} = B^{L}_{i,t}\text{.}$$

And for investments in capital

$$LL^{K}_{i,t,A} = -A^{K}_{i,t}\text{, } UL^{K}_{i,t,A} = A^{K}_{i,t}\text{, } LL^{K}_{i,t,B} = -B^{K}_{i,t} \text{ and } UL^{K}_{i,t,B} = B^{K}_{i,t}\text{.}$$

Where, as before

$$A_{i,t}^L = \sqrt{\frac{2b^L\alpha^L}{L_{i,t}}},\ A_{i,t}^K = \sqrt{\frac{2b^K\alpha^K}{K_{i,t}}},\ B_{i,t}^L = \sqrt{\frac{2b^L\alpha^L + \alpha^{KL}}{L_{i,t}}},\ B_{i,t}^K = \sqrt{\frac{2b^K\alpha^K + \alpha^{KL}}{K_{i,t}}}.$$

Now that we are maximizing the likelihood function with respect to four different threshold values, it is obvious that this is not possible while identifying five different cost parameters. Any combination of thresholds A^L, A^K, B^L and B^K can be derived from an infinite number of combinations of $b^L, b^K, \alpha^L, \alpha^K$ and α^{KL}.

Since the main interest of an empirical study where one would apply the model derived here lies the fixed cost parameters α^L, α^K and α^{KL}, and their relative sizes, an estimation of the relative size of b^L and b^K is satisfactory in this context. This can be used to limit the number of cost parameters to equal the number of threshold values, thus enabling exact identification. A satisfactory parameterization can therefore be applied by normalizing A^L, A^K, B^L and B^K with respect to $\sqrt{2b^L}$. This gives threshold expressions:

$$\tilde{A}_{i,t}^L \equiv \sqrt{\frac{\alpha^L}{L_{i,t}}}, \quad \tilde{A}_{i,t}^K \equiv \sqrt{\frac{\frac{b^K}{b^L}\tilde{\alpha}^K}{K_{i,t}}}, \quad \tilde{B}_{i,t}^L \equiv \sqrt{\frac{\tilde{\alpha}^L + \tilde{\alpha}^{KL}}{L_{i,t}}}, \quad \tilde{B}_{i,t}^K \equiv \sqrt{\frac{\frac{b^K}{b^L}\left(\tilde{\alpha}^K + \tilde{\alpha}^{KL}\right)}{K_{i,t}}}$$

[4.27]

Where a tilde indicates a parameter normalized with respect to $\sqrt{2b^L}$. Defining a variable $bb = \frac{b^K}{b^L}$ thus enables us to maximize the logarithmic likelihood function by setting threshold levels given by a unique combination of parameters bb, $\tilde{\alpha}^L$, $\tilde{\alpha}^K$ and $\tilde{\alpha}^{KL}$. This is particularly convenient because our main interest lies in the relative sizes of α^L, α^K and α^{KL}. One must however, keep in mind that individual estimates of alphas in the normalized thresholds do not directly translate back to their interpretation from equations [3.11] - [3.16].

Since we are mainly interested in relative sizes we can use these normalized expressions in the estimation. We must however assume that $b^L > 0$ to ensure that our estimated model has the right interpretation. Firstly, we can clearly not allow for a zero value. Second, because we need to specify investment regimes conditional on the sign of α^{KL} which in practice needs to be conditioned on the sign of $\tilde{\alpha}^{KL}$, we must add the assumption of $b^L > 0$. This should be a reasonable assumption remembering that b^L in the adjustment cost function

$$\begin{aligned} C(I_t, K_t, H_t, L_t) = & \left[p_t^{I+} I_t \cdot \mathrm{I}(I_t > 0) + p_t^{I-} I_t \cdot \mathrm{I}(I_t < 0) + \alpha^K + \frac{b^K}{2}\left(\frac{I_t}{K_t}\right)^2 \cdot K_t \right] \cdot \mathrm{I}(I_t \neq 0) \\ & + \left[p_t^{H+} H_t \cdot \mathrm{I}(H_t > 0) + p_t^{H-} H_t \cdot \mathrm{I}(H_t < 0) + \alpha^L + \frac{b^L}{2}\left(\frac{H_t}{L_t}\right)^2 \cdot L_t \right] \cdot \mathrm{I}(H_t \neq 0) \\ & + \alpha^{KL} \cdot \mathrm{I}(I_t \neq 0) \cdot \mathrm{I}(H_t \neq 0) \end{aligned} \quad [3.3]$$

is the convex cost parameter of labor adjustment, and therefore assumed to be positive.

4.6. Areas of Discontinuity of Derivatives

Although we at this point should be able find a unique combination of parameters to maximize our likelihood function, convergence is not necessarily easy when running the estimation procedure. Apart from setting feasible initial values, another problem easily arises as the software iterates to find a maximum. If either B^L or B^K is likely to be zero, the estimator will have a strong tendency to arrive at values of $\tilde{a}^L$, $\tilde{a}^K$ and $\tilde{a}^{KL}$ so that $\tilde{a}^{KL} = -\tilde{a}^L$ or $\tilde{a}^{KL} = -\tilde{a}^K$. This will indeed give zero values of B^L or B^K respectively. However, such values make it impossible for the software to calculate numerical derivatives (at $\sqrt{0}$), and consequently the estimation will break down.

In their theoretical model, Letterie et al. predicts that if $\alpha^K + \alpha^{KL} \leq 0$, this would mean that it will always be optimal for the firm to change the level of capital in every period it changes the level of employment. This is true because the effect of the fixed term α^K is completely balanced by the cost advantage represented by $\alpha^{KL} < 0$. This might indeed be the effect that is causing problems for the issue of convergence, and running simulations with different characteristics indicates that the estimation is more likely to break down if adjustment in one of the factors is always followed by adjustment of the other.

Since we may run into this issue when applying the estimation procedure to real life data, it might in practice be necessary to reformulate the likelihood function, in particular the threshold values, to avoid "unfriendly" values of α^{L}, α^{K} and α^{KL}. This however, can not be done without changing the formulation of the adjustment cost function in the theoretical model. The following section shows how such a reformulation can be carried out, and how new adjustment thresholds are derived.

4.7. Re-formulation of the Adjustment Cost Function [4]

4.7.1. Sequential adjustment

As before the management maximizes the value of the firm denoted by V(.) in equation (2.4.1). However, the new specification of the firms adjustment cost function is as follows:

$$
\begin{aligned}
C(I_t,K_t,H_t,L_t) = & \left[p_t^{I+} I_t \cdot \mathrm{I}(I_t > 0) + p_t^{I-} I_t \cdot \mathrm{I}(I_t < 0) + \alpha^K \boldsymbol{K_t} + \frac{b^K}{2}\left(\frac{I_t}{K_t}\right)^2 \cdot K_t \right] \cdot \mathrm{I}(I_t \neq 0) + \\
& + \left[p_t^{H+} H_t \cdot \mathrm{I}(H_t > 0) + p_t^{H-} H_t \cdot \mathrm{I}(H_t < 0) + \alpha^L \boldsymbol{L_t} + \frac{b^L}{2}\left(\frac{H_t}{L_t}\right)^2 \cdot L_t \right] \cdot \mathrm{I}(H_t \neq 0) + \dots \\
& + \alpha^{KL} \sqrt{\boldsymbol{K_t L_t}} \cdot \mathrm{I}(I_t \neq 0) \cdot \mathrm{I}(H_t \neq 0)
\end{aligned}
\qquad [4.28]
$$

[4] The derivation in this subsection follow the derivation of optimal factor adjustment levels in Letterie et. al. (2009)

where the changes from the original formulation are given in bold letters. Under this specification, we let fixed cost components depend on firm size. The fixed costs of capital adjustment grow with the level of capital, K_t, and the fixed costs of labor adjustment grow with the number of workers employed by the firm, L_t. The size of a possible interrelation cost is dependent on the levels of both capital and labor, through the term $\sqrt{K_t L_t}$. As before capital and labor evolve according to equations [3.4] and [3.5] respectively. Maximization of firm value yield identical first-order conditions as were shown in equations [3.6] and [3.7], and optimal rates of investments and hiring as in equations [3.8] and [3.9].

The management of the firm will wish to adjust the level of capital if the benefits associated with an adjustment exceed the costs. That is if

$$\left(\lambda_t^K - p_t^I\right) I_t > \alpha_0^K K_t + \frac{b^K}{2}\left(\frac{I_t}{K_t}\right)^2 K_t \qquad [4.29]$$

which gives

$$\frac{1}{2b^K}(\lambda_t^K - p_t^I) \geq \alpha^K \qquad [4.30]$$

Now we have the following threshold for capital adjustment:[5]

$$\left|\lambda_t^L - p_t^I\right| > \sqrt{2b^K \alpha^K} \equiv A^{K*} \qquad [4.31]$$

And similarly we have the following threshold for changes in labor:

$$\left|\lambda_t^L - p_t^H\right| > \sqrt{2b^L \alpha^L} \equiv A^{L*} \qquad [4.32]$$

[5] Asterisk denotes factor adjustment thresholds after re-formulation of the adjustment cost function.

4.6.2. Simultaneous adjustment

In the case of simultaneous adjustment, the cost component α^{KL} will come into play, and we can derive thresholds that are valid under the new parameterization of the adjustment cost function. As before, it is optimal to adjust an additional factor of production if the net benefits associated with that adjustment exceed the fixed costs of that second input (α^{K} or α^{L}) plus the cost of interrelation $\alpha^{KL} > 0$. For an additional investment in capital, this gives

$$\left(\lambda_t^K - p_t^I\right) I_t > \alpha_0^K K_t + \frac{b^K}{2}\left(\frac{I_t}{K_t}\right)^2 K_t + \alpha^{KL}\sqrt{K_t L_t} \qquad [4.33]$$

, which can be rewritten as

$$\frac{1}{2b^K}\left(\lambda_t^K - p_t^I\right)^2 K_t > \alpha^K + \frac{\alpha^{KL}}{K_t}\sqrt{K_t L_t} \qquad [4.34]$$

In the case of an additional adjustment in the number of employees this translates into

$$\frac{1}{2b^L}\left(\lambda_t^L - p_t^H\right)^2 L_t > \alpha^L + \frac{\alpha^{KL}}{L_t}\sqrt{K_t L_t} \qquad [4.35]$$

We now have the following two thresholds for simultaneous adjustments:

$$\left|\lambda_t^K - p_t^I\right| > \sqrt{2b^K\left(\alpha^K + \alpha^{KL}\sqrt{\frac{L_t}{K_t}}\right)} \equiv B^{K*} \qquad [4.36]$$

$$\left|\lambda_t^L - p_t^H\right| > \sqrt{2b^L\left(\alpha^L + \alpha^{KL}\sqrt{\frac{K_t}{L_t}}\right)} \equiv B^{L*} \qquad [4.37]$$

As before, we can estimate the model with normalized threshold expressions and maximize the likelihood function with respect to parameters bb, $\tilde{\alpha}^L$, $\tilde{\alpha}^K$ and $\tilde{\alpha}^{KL}$.

When applying these new thresholds to the log likelihood function, we avoid the problems using the original parameterization, described in section 4.3. The reason for this can clearly be seen as the term in the brackets on the right hand side of equations[4.36] and [4.37] no longer only consist of a sum of $\alpha^K + \alpha^{KL}$ and $\alpha^L + \alpha^{KL}$ respectively. The last term is now a product of α^{KL} and the square root of the ratio of labor to capital, or capital to labor. This means that a solution where B^L or B^K is set to zero for all observations can not be accomplished by simply setting α^L or α^K equal to $-\alpha^{KL}$, but must result from $\alpha^L = \alpha^{KL} = 0$ or $\alpha^K = \alpha^{KL} = 0$.

On a further note, one of the predictions of Letterie et al. was that the importance of possible interrelations is dependent upon firm size, represented by L_t and K_t. More precisely the prediction was that firm size would decrease the importance of the interrelation term α^{KL} through the denominators L_t and K_t in the investment thresholds B_I^L and B_I^K respectively. Under the new specification of the cost function, we lose the inclusion of these denominators, so that the roles of L_t and K_t become slightly different.

A large L_t will still undermine the importance of α^{KL} in labor adjustments, and K_t will do the same for capital adjustments. However, under the alternative specification, the relative levels and not the absolute levels of production factors are decisive. A large labor-to-capital ratio (number of workers for every machine) will decrease the importance of α^{KL} in decisions of hiring when capital is adjusted in the same period. Likewise, it will increase the importance of α^{KL} in decisions of capital investment when labor is adjusted in the same period. In other words, a labor-intensive producer will be relatively unconcerned about whether or not the firm also invests in capital when deciding on changes in employment. Interrelations in hiring decisions should, on the other hand, matter more when it comes to a capital-intensive producer.

5. Data

To be able to observe the performance of the Maximum Likelihood estimator, a simulated panel dataset is constructed for use in an illustrative analysis. In this section, some characteristics of data in similar empirical studies are discussed. Summary statistics for data simulated to test our derived model follows. The simulated data is created to replicate moments seen in datasets used to analyze firms' factor demand.

5.1. Characteristics of Data in Empirical Research

In studies concerning factor adjustments of labor, capital or both, we can find some general characteristics that seem to be relatively consistent in the datasets applied. We can use some of these general qualities as a basis for our simulations to ensure that we are not drifting to far off from the real world.

For our purpose, the feature of main interest lies in the variables of hiring and investment. While zero adjustment of capital make up about eight per cent of the observations from US manufacturing plants used by Cooper and Haltiwanger (2005), Norwegian data used in a study by Nilsen and Schiantarelli (1998) contain about twenty-one per cent inaction observations. Rota (2004), studying adjustments of labor, apply a dataset where zero adjustment stand for around twenty per cent of observations of Italian firms over a ten year period. Contreras (2006) studies simultaneous adjustments of labor and capital for Colombian plants, and inaction in investments make up about nineteen per cent of the observations, while inaction in labor adjustment stand for about thirteen per cent.

Further, investments in capital equipment seem on average to be positive, while the average of hiring or firing seems to lie closer to zero. Letterie et al. (2004) use data from the Dutch manufacturing sector with an average labor adjustment ratio of 0.015 and an average investment ratio of 0.098. Similarly, the average investment rate in the data applied by Cooper and Haltiwanger (2005) is about twelve per cent and the fraction of observations with negative investments only ten per cent.

Also, most data used indicates positive correlation between hiring and investment decisions, although of different magnitude. In a working paper by Narazzani (2009), firm level correlation seems substantial, while the data studied by Contreras (2006) only inhabit a correlation of 0.057.

5.2. Simulation

Variables are generated for 1000 observations to represent hiring (H/L) and investment rates (I/K) together with levels of labor (L) and capital (K) and variables to represent shadow values net of purchase/resale prices for the two inputs (q^K and q^L). We establish positive correlation between shadow values and their respective adjustment rates, but also set a low positive correlation between shadow values and adjustments of the second input. In addition to positive correlation between the factor adjustment rates, we simulate a positive correlation between the two marginal values. Investment rates are constructed to be slightly higher than hiring rates on average, and the same is done for the average level of capital compared to the average level of labor.

We also construct year variables, so that we work with a balanced panel data set of 200 firms over five years. However, we do not generate any time specific or firm specific effects, so that any such effects are purely random

To give the data qualities similar to those seen in studies of investment behavior, we take a few extra measures in the generating process. Firstly we drop observations with values of labor (L) and capital (K) smaller than 25 to avoid negative and small values. Secondly, we wish to observe the same accumulation around zero adjustments as is seen in a lot of firm level data. This is done by letting all hiring and investment rates with an absolute value less than 0.05 take a zero value. The effect is clearly seen from the histograms in figures (5.1) and (5.2) as they display apparent peaks around the zero values in what otherwise looks like bell shaped distributions.

TABLE 5.1: CORRELATION MATRIX

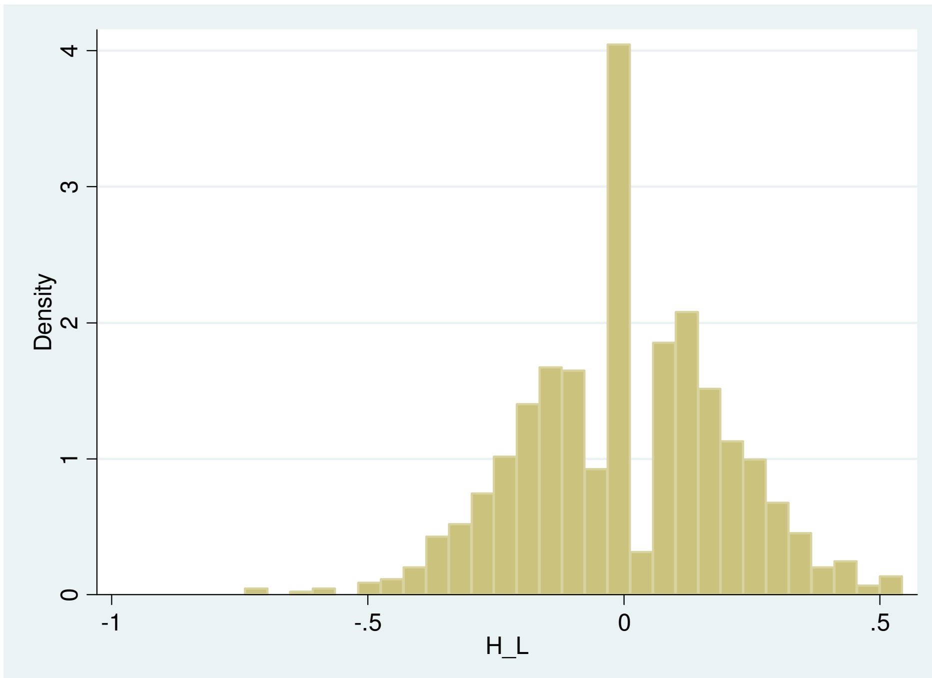
	H/L	I/K	L	K	q^K	q^L
H/L	1					
I/K	0.7238	1				
L	-0.0095	0.0295	1			
K	-0.0700	-0.0539	0.5778	1		
q^K	0.2198	0.5595	0.0241	-0.0383	1	
q^L	0.5604	0.1673	-0.0329	-0.0633	0.5066	1

TABLE 5.2: SUMMARY STATISTICS

Variable	Obs.	Mean	Std. dev.	Min.	Max.
H/L	1000	.003465	.198821	-.7398592	.5436076
I/K	1000	.0105616	.2002771	-.5664857	.7501802
L	1000	358.0854	240.7971	25.46533	1389.317
K	1000	7884.791	3779.305	89.18464	20929.77
q^K	1000	3.084015	2.986225	-8.111652	12.74151
q^L	1000	3.916717	4.16364	-9.872766	20.57577

When we look at the joint distribution of the discrete variables of hiring and investment in table 5.3, we see that the simulated sample is dominated by positive investment in capital, but also that observations with a simultaneous positive adjustment have the highest count. Hiring seems only to occur along with positive investments, while firing is relatively evenly distributed over investments in capital. Negative investments in capital are almost exclusively occurring alongside firing, while positive investments are more evenly distributed.

FIGURE 5.1: HIRING



When it comes to inaction, we see a very small amount of simultaneous inaction, which to the eye indicates a relatively small role played by non-convexities, even though the indications from the individual histograms point in another direction. Given the possibility of interrelations in the adjustment cost functions and lessons from earlier research, we also know that we should be careful about which conclusions we draw from simply investigating descriptive statistics of investment patterns. Perhaps these distribution diagrams show us just that. The next section will present the estimation results after applying our empirical model on the simulated data. The hope is that these results can shed some light on possible underlying factors.

FIGURE 5.2: INVESTMENT

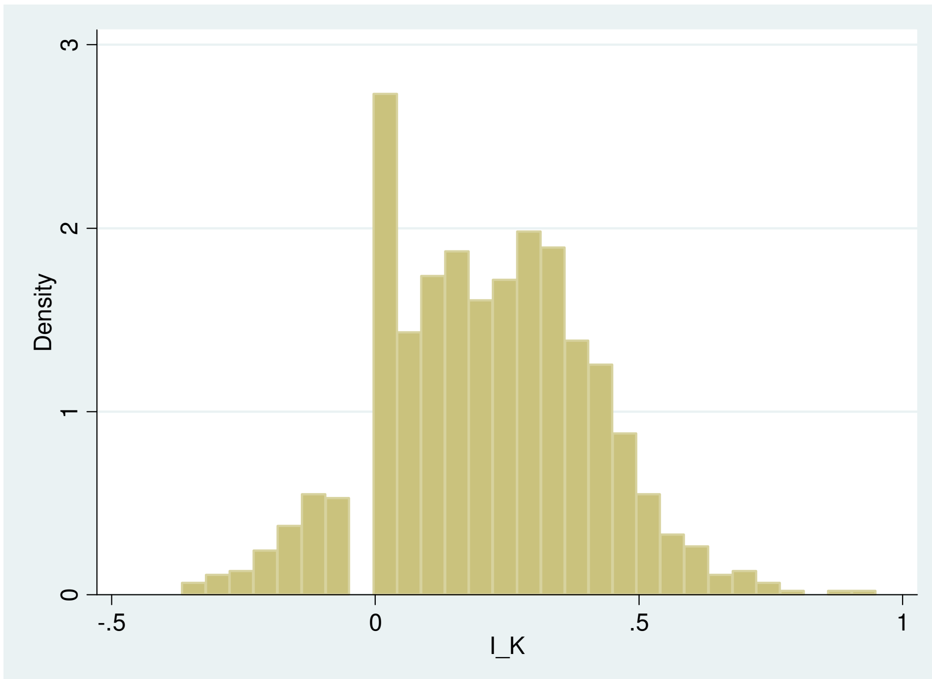


TABLE 5.3: HIRING & INVESTMENT (JOINT DISTRIBUTION)

	Hiring			
Invest	-1	0	1	**Total**
-1	87	3	1	91
0	92	24	8	124
1	214	152	419	785
Total	393	179	428	1000

It is noteworthy how different ways of generating zeros in the two adjustment processes give different implications for what we would expect of the role of an interrelation cost parameter α^{KL}. Working on generating data in does in some ways give insight to how difficult it can be to spot an interrelation effect from the cost function just by an eyeball test of descriptive statistics. In appendix B, descriptive statistics of an alternative simulation can be found. The correlation matrix and summary statistics of the two datasets applied look very much similar. However, these matrices are not really the main determinants of observations of joint inaction. The main difference between the two simulations presented lies in the process of generating extra observations of zero adjustments.

The first dataset is given extra zeroes by letting observations of the variables H/L and I/K take a zero value if the absolute value is less than 0.05. The approach is altered when generating the second dataset as we want to see the effects of a cost reduction from simultaneous adjustments. We increase the amount of observations with no adjustments to increase the probability of significant nonconvex parameters, and zeroes are generated for absolute values less than 0.10. This creates both individual and joint zeroes dependant on the correlation between H/L and I/K. In addition, extra zero observations are generated for all observations where one of the adjustment variables has an absolute value less than 0.20, but only if the other adjustment variable is a zero. The intention behind this is to let those observations that are simultaneously larger than the first threshold have “an advantage” over those that only contain one adjustment variable above that particular limit, thus simulating a simultaneous cost advantage. In sum, we would therefore expect interrelations to play a more prominent role in the behavior simulated in the appendix, as no such trick is applied in the simulated data presented above.

6. Estimation Results

6.1. Simultaneous Estimation

Table 6.1 presents estimation results for two different simulated data sets with standard deviations in parentheses. Column (i) contains estimates for the data presented in chapter 5, while column (ii) contains estimates for a dataset with characteristics presented in appendix B. Our parameterization using a reformulated adjustment cost function allows estimation of coefficients of simulated marginal q's from equations [4.4] and [4.5], normalized with respect to $\sqrt{2b^L}$ (denoted $\tilde{\beta}_1^L$ and $\tilde{\beta}_1^K$). We also obtain normalized coefficient estimates of constant terms which include $\left(\beta_0^K + p_{i,t}^I\right)$ and $\left(\beta_0^L + p_{i,t}^H\right)$ from the same equations, in addition to year dummies, though they are not reported here.

Further, estimates are reported for logarithmic values of the normalized fixed adjustment cost parameters, $\tilde{\alpha}^L$ and $\tilde{\alpha}^K$, and the relative sizes of convex cost parameters, $\sqrt{b^K/b^L}$. These are on logarithmic form to restrict estimation to positive values. No such restriction is applied to the estimation of the interrelation cost parameter $\tilde{\alpha}^{KL}$, although we must remember that this of course is normalized with respect to $\sqrt{2b^L}$ as indicated by a tilde.

The last estimate is of $\tilde{\rho}$, which we remember is the correlation coefficient in the bivariate normal distribution function in equation [4.14]. A large positive estimate of this parameter indicates a high level of positive correlation between the error terms $\varepsilon_{i,t}^L$ and $\varepsilon_{i,t}^L$ in the latent equations [4.4] and [4.5]. This would mean that any unobserved factors affecting the profitability of one input have a tendency of affecting the profitability of the second input in the same direction. Obviously, the opposite would be true in the case of a large negative estimate of $\tilde{\rho}$.

Table 6.2 presents derived parameters and ratios that we are able to obtain from the results in table 6.1. P-values after nonlinear tests are reported in the parentheses. These are perhaps the results of main interest, since normalized parameter estimates individually are of limited use. They include as well as comparisons of convex and nonconvex cost parameter sizes for labor and capital, the relative sizes of fixed costs compared to the estimated interrelation component and a comparison of estimated total threshold sizes, A^L / A^K .

From column (i) in table 6.1 we see that although only two of the parameter estimates are statistically significant, the results indicate positive relationships between marginal q's and factor profitability through positive estimates of $\tilde{\beta}_1^L$ and $\tilde{\beta}_1^K$. This is certainly something we would expect from real life firm behavior. The estimated correlation coefficient $\tilde{\rho}$ is significant and indicates a strong negative correlation between the error terms $\varepsilon_{i,t}^L$ and $\varepsilon_{i,t}^L$. This suggests that unobserved effects that influence factor profitability seem to have opposite effects for the two production inputs. The estimates of log transformed fixed cost parameters $\tilde{\alpha}^L$ and $\tilde{\alpha}^K$, indicate a relatively large fixed cost component in labor adjustment costs and a relatively small fixed cost component for capital. The opposite is true for the estimate of the relative sizes of convex cost parameters ($\ln\sqrt{b^K/b^L}$) where the indication points towards a relatively large convex cost component for capital compared to that of labor adjustment. The indications are supported by the post-estimation results presented in table 6.2, where relative sizes of cost parameters and threshold levels are shown explicitly.

Although none of the estimates are statistically significant, we note that they are perhaps in contrast to what we would expect from observed firm level behavior regarding the degree of fixity of labor and capital. Since capital is usually considered as a less flexible input than labor, we would for a real data set, expect the fixed component of adjustment costs to be large relative to that of labor, and perhaps, but not necessarily, the opposite relationship for convex costs. It would at least be likely that the threshold size would be larger for capital than labor, given that capital adjustment in most cases is more lumpy than labor adjustment.

TABLE 6.1: PARAMETER ESTIMATES
(Reformulated adjustment cost function)

	(i)	(ii)
$\tilde{\beta}_1^L$	.01606 (.01157)	.09757 (.01383)***
$\tilde{\beta}_1^K$	.08160 (.02223)***	.28485 (.02754)***
$\ln\left(\tilde{\alpha}^L\right)$	3.5222 (3.1875)	4.5280 (.40275)***
$\ln\left(\tilde{\alpha}^K\right)$	-4.2445 (7.7759)	1.2353 (.34863)***
$\ln\sqrt{b^K/b^L}$	4.1122 (7.7635)	-1.0286 (.37800)***
$\tilde{\alpha}^{KL}$	.00053 (.00826)	-7.9402 (4.6760)*
$\tilde{\rho}$	-.84752 (.0724482)***	.15809 (.10708)
N	1000	1000
Log L	-658.49	-951.53

Notes: Standard deviations are given in parentheses. * Indicates significance at 10% level. ** Indicates significance at 5% level. *** Indicates significance at 1% level. Column (i) presents results from estimation on the dataset presented in section 5. Column (ii) presents results from estimation on a dataset created to generate $\alpha^{KL} < 0$. $\hat{\rho}$ is the estimated correlation coefficient for the error terms $\varepsilon_{i,t}^L$ and $\varepsilon_{i,t}^K$ in the latent equations.

Column (ii) of table 6.1 displays estimation results for the second set of simulated data. First, we notice how nearly all coefficient estimates are statistically significant. Again we find positive relationships between marginal q's and factor profitability through positive estimates of $\tilde{\beta}_1^L$ and $\tilde{\beta}_1^K$. The estimate of the correlation coefficient however, has gone from negative to not significantly different from zero. As in column (i), the estimate of fixed cost parameters $\tilde{\alpha}^L$ and $\tilde{\alpha}^K$ indicate a relatively large fixed cost component for adjustments in labor, which is confirmed by the post-estimate of $\sqrt{\alpha^L/\alpha^K}$ in table 6.2.

The estimation of the relative sizes of convex cost parameters also strongly indicates a relatively large parameter for labor. In contrast to the results in column (i), the relative differences are statistically significant, and so is the difference in threshold levels, measured by A^L / A^K.

TABLE 6.2: DERIVED PARAMETERS AND RATIOS
(Reformulated adjustment cost function)

	(i)	(ii)
$\sqrt{\tilde{\alpha}^L}$	33.859 (0.754)*	92.578 (0.013)*
$\sqrt{\tilde{\alpha}^K}$	.01434 (0.898)*	3.4393 (0.004)*
$\sqrt{\alpha^L / \alpha^K}$	2360.8 (0.907)**	26.918 (0.006)**
$\sqrt{\alpha^L} / \alpha^{KL}$	63514 (0.950)**	-11.659 (0.011)**
$\sqrt{\alpha^K} / \alpha^{KL}$	26.903 (0.903)**	-.43315 (0.000)**
$\sqrt{b^K / b^L}$	61.082 (0.899)**	.35748 (0.000)**
A^L / A^K	38.650 (0.773)**	75.299 (0.0185)**

Notes: The table presents post-estimations of derived ratios together with p-values in parentheses after non-linear tests. * Indicates a p-value of a test with a null hypothesis of size = 0. ** Indicates a p-value of a test with a null hypothesis of size = 1. Column (i) presents results from estimation on the dataset presented in section 5. Column (ii) presents results from estimation on a dataset created to generate $\alpha^{KL} < 0$.

In addition to the obvious result that more estimates turn out to be significant in column (ii), the estimate of the interrelation cost parameter, $\tilde{\alpha}^{KL}$, has gone from very close to zero in the first column to a borderline case of a significantly negative estimate. In addition, the post-estimation of $\sqrt{\alpha^L} / \alpha^{KL}$ and $\sqrt{\alpha^K} / \alpha^{KL}$ tell us that the interrelation parameter is relatively small compared to α^L, and relatively large compared to α^K. A small relative size in absolute terms seems most reasonable, given that possible interrelation effects in most studies appear to be subtle.

The estimation presented in column (ii) is applied to data generated to inhabit strong signs of both nonconvexities and interrelations. The results presented indicate that our routine is equipped to pick up such differences in data characteristics. However, it is crucial to include a discussion of what the effect the estimate of $\tilde{\alpha}^{KL}$ may have on the identification of the remaining parameters. In table 6.1 we see a striking difference between the estimates of $\tilde{\alpha}^{K}$ and the relative sizes of b^K and b^L. In column (i) a large estimate of $\ln\sqrt{b^K/b^L}$ goes hand in hand with a small estimate of $\ln\left(\tilde{\alpha}^K\right)$. This happens while standard deviations indicate very uncertain estimates. This might not look too dramatic, but may nevertheless lead us to investigate the possible role of $\tilde{\alpha}^{KL}$.

Perhaps the most important motivational factor behind estimating a likelihood function described in this section, is to allow for interrelations in our adjustment cost function to see how this affects decisions of factor adjustment. We have derived an empirical model that allows us to test the null hypothesis of $\alpha^{KL}=0$. Given that we want to test this hypothesis, we should also consider what happens to our model if the true value of the interaction term is in fact zero.

No interaction term would, as we remember make our adjustment thresholds equal for sequential and simultaneous adjustments. That is

$$A_{i,t}^{L}=B_{i,t}^{L}=\sqrt{\frac{2b^{L}\alpha^{L}}{L_{i,t}}} \quad \text{and} \quad A_{i,t}^{K}=B_{i,t}^{K}=\sqrt{\frac{2b^{K}\alpha^{K}}{K_{i,t}}}$$

Or given by our four parameters used in estimation

$$\tilde{A}_{i,t}^{L} = \tilde{B}_{i,t}^{L} = \sqrt{\frac{\tilde{\alpha}^{L}}{L_{i,t}}} \quad \text{and} \quad \tilde{A}_{i,t}^{K} = \tilde{B}_{i,t}^{K} = \sqrt{\frac{\frac{b^{K}}{b^{L}}\tilde{\alpha}^{K}}{K_{i,t}}}$$

The equation above tells us that under the condition of no interrelation term, estimation of our previously chosen parameters will make our model unidentifiable. The problem arises because one threshold level for capital adjustment in our likelihood function now can be expressed through an infinite number of combinations of parameters b^K/b^L and $\tilde{\alpha}^K$. This is certainly problematic since one of the key questions in this analysis is how we can estimate the size of α^{KL}. Now that we see how our estimation approach will break down if a null hypothesis of $\alpha^{KL} = 0$ is true, it should make us question the attractiveness of the log likelihood function for our purpose. A likely outcome of any empirical application is that we end up with an estimated $\tilde{\alpha}^{KL}$ that differs insignificantly from zero. If this would be the case, estimates of b^K/b^L and $\tilde{\alpha}^K$ would as a consequence be highly questionable, and we would possibly be unable to identify the size of convex costs compared to nonconvex costs in addition to the relative size of convex cost parameters b^K/b^L and nonconvex parameters α^L/α^K.

In an empirical application we would roughly stated, as the interrelation term moves towards zero, expect the likelihood function to move closer to an area where the parameters in question are unidentifiable. In such a case we would expect the derivatives that describe the marginal effects on the likelihood value of adjusting parameter values to become very small as they move towards zero. Accordingly we may observe very large or very small estimates of either one of the estimates, even though none of them are significantly different from zero. The intuition behind this stems from what we know about estimation by Maximum Likelihood, where standard deviations are estimated on the basis of second derivatives. A relatively small standard deviation which indicates a reliable coefficient estimate corresponds to a large negative second derivative and vice versa. In the opposite case, adjusting the coefficient estimate marginally is not important for the likelihood value, resulting in an uncertain estimate.

This potential identification problem may be a plausible explanation for some of the differences between column (i) and (ii). In any case, the possibility that the parameters in our model are not identifiable under certain circumstances is of great concern. Therefore, in future work on an empirical application of this model, this weakness should be addressed. For that purpose, a new formulation of the adjustment cost function might be necessary.

6.2 Separate Estimation

For comparison purposes, we also estimate two separate ordered probit models by a customized ML approach. The difference is that this time we assume that the interrelation term, $\alpha^{KL} = 0$, and that we assume the two adjustment decision equations are completely unrelated, so that there exists no correlation between the error terms in the two equations $(\rho = 0)$. Table 6.2 displays the results of the separate estimations. We notice how we find the same relationship between the simulated marginal q's and factor adjustment as before. As expected, coefficients β_{q^L} and β_{q^K} are positive, which means that they have a positive effect on the latent variables $\lambda_{i,t}^{L}$ and $\lambda_{i,t}^{K}$ respectively. In a traditional limited dependent variable language, we interpret this as a positive effect of the marginal value of one input factor on the expected utility of one additional unit of that input. We further note that the threshold for capital adjustment is estimated to be relatively large compared to that of labor. Even though the difference seems small and the threshold estimates after simultaneous estimation were not significantly different, this is the opposite of what was indicated in table 6.1 using the same data.

TABLE 6.2: PARAMETER ESTIMATES -

SEPARATE ESTIMATION

	Labor	**Capital**
$\tilde{\beta}_1^L$	.16230 (.01073)***	-
$\tilde{\beta}_1^K$	-	.23767 (.01859)***
$\ln\left(A_S^L\right)$	-1.2983 (.06832)***	-
$\ln\left(A_S^K\right)$		-1.102545 (.08239)***
N	1000	1000
Log L	-906.51	-565.42

Notes: *** Indicates estimates significant at 1% level. Separate likelihood functions only allow us to estimate one threshold parameter for each factor. These are denoted A_S^L for labor, and A_S^K for capital. To restrict the estimation to positive values, we estimate log transformed variables. The relative size of A_S^L / A_S^K is approximately 0.82.

7. Concluding Remarks

In this analysis we have derived and discussed the properties of an estimable model of factor adjustment that allows for nonconvex adjustment costs. In addition our model allows for the existence of an interrelation cost that may reduce or increase the costs when factor adjustments in two inputs occur simultaneously. Further, we have discussed parameterization necessary to identify our parameters of main interest. In sum, this text shows how a likelihood function can be written to estimate adjustment cost parameters that can be traced directly back to a theoretical framework for adjustment costs and factor demand.

The model is tested on simulated panel data which replicate moments often seen in datasets used to analyze firms' factor demand. Two different datasets are created to investigate the properties of the model. The datasets have characteristics that are similar in most respects, but are purposely created to differ in the patterns of zero adjustments. Running the estimation routine, we are able to obtain estimates of nonconvex adjustment cost parameters. The estimations on simulated data show how the ML estimates differ between a case with an interrelation cost parameter close to zero and a case where we estimate the parameter to be negative.

While presenting the performance of our customized Maximum Likelihood estimation, the text has also pointed out crucial assumptions and weaknesses to such an approach. The simultaneous estimation allowing for correlation between factor demand equations is obviously a strength. Empirical findings strongly indicate that ignoring correlation could lead to biased estimates of adjustment costs. Furthermore, working on simulating and estimating interrelation costs leaves the impression that spotting the existence of such costs by viewing the data separately is extremely difficult. However, for the ML estimator applied in this thesis, identification problems arising in the existence of an interrelation parameter equal to zero is clearly a weakness that should be dealt with in future work.

As a natural extension of the work on this thesis, the model should of course be tested on real data. Firstly by applying the likelihood function derived in this thesis. As we have spotted obvious room for improvement, a more robust parameterization should also be found, most likely together with a reformulation of the underlying adjustment cost function itself.

Considering that data on factor adjustment generally contains information about the actual size of the factor adjustment, it would be desirable to apply an estimation technique that can make use of this information. In a probit estimation like the one derived in this thesis we do not make use of all available information. Properties of a Tobit-like likelihood function to obtain more efficient estimates could therefore be worth exploring.

References

(i) Estimating models with limited dependent variables:

Cameron, A. Colin and Pravin K. Trivedi (2005): Microeconometrics: methods and applications, Cambridge University Press, Cambridge.

Greene, William H. (2008): Econometric Analysis, 6th edition, Pearson Prentice Hall, Upper Saddle River, N. J.

Verbeek, Marno (2008): A guide to modern econometrics, 3rd edition, Chichester: Wiley.

Wooldridge, Jeffrey M. (2006): Introductory econometrics: a modern approach, 3rd edition, Thomson South-Western, Mason, Ohio.

Gould, W., J. Pitblado and W. Sribney (2006): Maximum Likelihood Estimation with Stata, Stata Press, College Station.

(ii) Factor demand, adjustment costs and uncertainty

Abel, A.B. and J.C. Eberly (1994): "A unified model of investment under uncertainty". (*American Economic Review*, 84, 1369-1384.)

Abel, A.B. and J.C. Eberly (1998): "The Mix and Scale of Factors with Irreversibility and Fixed Costs of Investment". (*Carnegie-Rochester Conference Series on Public Policy*, 48, 101-135.)

Bloom, N. (2009): “The impact of uncertainty shocks”, (*Econometrica*, 77, 623-685.)

Blundell, R., S. Bond, M. Devereux, F. Sciantarelli (1992): “Investment and Tobin’s Q: Evidence from Company Panel Data”. (*Journal of Econometrics*, 51, 233-257.)

Caballero, R. J. and E. M. R. A. Engel (1999): “Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach”. (*Econometrica*, 67, 783-826.)

Contreras, J.M. (2006): “Essays on factor adjustment dynamics”. (PhD Thesis, University of Maryland, Department of Economics.)

Cooper, R. W. and J. C. Haltiwanger (2006): “On the Nature of Capital Adjustment Costs”. (*Review of Economic Studies*, 73, 611-633.)

Dixit, A. (1997) “Investment and employment dynamics in the short and the long run”. (*Oxford Economic Papers*, 49, 1-20.)

Eberly, J.C. and J.A. van Mieghem (1997): “Multi-factor dynamic investment under uncertainty”. (*Journal of Economic Theory*, 75, 345-387.)

Galeotti, M. and F. Schiantarelli (1991): “Generalized Q Models for Investment”. (*Review of Economics and Statistics,* 73, 383-392.)

Hamermesh, D. S. (1996) Labor Demand, Princeton University Press, Princeton, N. J.

Hamermesh, D. S. and G. A. Pfann (1996): “Adjustment Costs in Factor Demand”. (*Journal of Economic Literature*, 34, 1264-1292.)

Jorgenson, D. W. (1963): “Capital Theory and Investment Behavior”. (*The American Economic Review*, 53, 247-259.)

Letterie, W., Ø.A. Nilsen and G. A. Pfann (2009): ”Interrelated Factor Demand with Nonconvex Adjustment Costs”. (*Working paper*)

Letterie, W. A., G. A. Pfann and J. M. Polder (2004): ”Factor adjustment spikes and interrelation: an empirical investigation”, (*Economic Letters*, 85, 145-150.)

Merz, M. and E. Yashiv (2007): “Labor and the value of the firm” (*American Economic Review*, 97, 1419-1431.)

Nadiri, M.I. and S. Rosen (1969): “Interrelated factor demand functions”. (*American Economic Review*, 59, 457-471.)

Narazzani, E. (2009): “Labour and Capital Adjustment Dynamics Using Bivariate Random Probit”. (*Part of PhD Thesis, University of Turin.*)

Nilsen, Ø.A, K.G. Salvanes and F. Schiantarelli (2007): “Employment changes, the structure of adjustment costs, and plant size”, (*European Economic Review*,,51, 577-591.)

Nilsen, Ø. A., A. Raknerud, M. Rybalka and T. Skjerpen (2009): “Lumpy Investments, Factor Adjustments and Labour Productivety”. (*Oxford Economic Papers*, 61, 104-127.)

Nilsen, Ø. A., F. Schiantarelli (2003): “Zeros and Lumps in Investments: Empirical Evidence on Irreversibilities and Non-convexities.” (*Review of Economics and Statistics*, 85, 1021-1037.)

Rota, P. (2004) “Estimating Labor Demand with Fixed Costs”. (*International Economic Review*, 45, 25-48.)

Sakellaris, P. (2004): “Patterns of plant adjustment”. (*Journal of Monetary Economics*, 51, 425-450.)

Shapiro, M. (1986): “The dynamics demand for capital and labor". (*Quarterly Journal of Economics,* 51, 513-542.)

Tobin, J. (1969): “A General Equilibrium Approach to Monetary Theory”. (*Journal of Money Credit and Banking*, 1, 15-29)

Appendix:

(A.1) Treating Marginal Values

In this thesis we have considered determinants of firms' investment behavior with a focus on non-convexities in adjustment cost functions. What should be seen as the most important factor of influence, namely the marginal values of capital and labor has to a large extent been left out of the discussion. In the theoretical framework, the variables are denoted by λ_t^L and λ_t^K, and can be traced back to the Lagrangean objective function applied to maximize firm value $V(.)$. These shadow values can be affected by changes in demand or technological changes and other determinants of factor productivity, like utilization rates for both factors.

Working with simulated data, we have assumed that the unobserved marginal values can be described by a linear relationship with marginal q's which we generate with desired characteristics into our dataset. We have simply assumed that this is a reasonable approach and gone on with the estimation. Since we mainly are interested in deriving an estimation technique, the assumption is not really critical for our purpose. However, since the thesis aims to present a way to estimate adjustment cost functions with interrelations applicable to empirical research, we should consider what actually lies behind this assumption, and what implications it might have for our ability to obtain valid empirical results.

The lack of observability simply results from the fact that marginal values are never reported, and the general way of dealing with this is to apply a suitable linear approximation. It can be shown analytically that the shadow values for capital and labor are given by

$$\lambda_t^K = E_t\left[\sum_{s=0}^{\infty}\frac{\left(1-\delta^K\right)^s}{\left(1+r\right)^{s+1}}\left(\frac{\partial F\left(A_{t+s+1},K_{t+s+1},L_{t+s+1}\right)}{\partial K_{t+s+1}}+\frac{b^K}{2}\left(\frac{I_{t+s+1}}{K_{t+s+1}}\right)^2\right)\right] \qquad [9.6]$$

$$\lambda_t^L = E_t\left[\sum_{s=0}^{\infty}\frac{\left(1-\delta^L\right)^s}{\left(1+r\right)^{s+1}}\left(\frac{\partial F\left(A_{t+s+1},K_{t+s+1},L_{t+s+1}\right)}{\partial L_{t+s+1}}+\frac{b^L}{2}\left(\frac{H_{t+s+1}}{L_{t+s+1}}\right)^2\right)\right] \qquad [9.7]$$

respectively.

These equations tell us that the marginal values are functions of expectations over an infinite number of discrete time periods *t*. Further, we clearly see that the marginal values of input factor adjustments can be separated into two components. The first component and perhaps the most obvious one, is the marginal product one production factor in all future periods. The second component, the term on the right side in the brackets, is perhaps less obvious. We can consider this a cost advantage of an additional unit of capital or labor in period *t*. This cost advantage arises because an additional unit today affects the expected level of adjustment costs tomorrow. Now the reasoning gets a little tricky. Since the shadow value in reality is an expectation, the investment decision in period *t* must clearly depend on the decision in every future period $t+s$. However, we also know that future investment decisions will be dependent on the firms' previous choices. How to deal with this is in no way clear cut, but under certain assumptions approximations can be derived.

In most of the literature on this subject, one assumes convex adjustment costs which implies smooth and relatively small adjustment rates. In section A.2 it is shown that if adjustments are small, shadow values can be approximated by:

$$\lambda_t^K \approx \left[\sum_{s=0}^{\infty}\frac{\left(1-\beta\right)\left(1-\delta^K\right)^{s(1-\beta)-\beta}}{\left(1+r\right)^{s+1}}\left(\frac{Y_t}{K_t}\right)\right] \propto \theta\left(\frac{Y_t}{K_t}\right) \qquad [9.15]$$

and

$$\lambda_t^L \approx \left[\sum_{s=0}^{\infty} \frac{(1-\beta)(1-\delta^L)^{s(1-\beta)-\beta}}{(1+r)^{s+1}} \left(\frac{Y_t}{L_t} \right) \right] \propto \theta \left(\frac{Y_t}{L_t} \right) \qquad [9.16]$$

, which expresses shadow values as functions of sales-to-capital and sales-to-labor ratios in period *t.* This approximation is fairly standard in empirical work, but implies an assumption that adjustments are small. When we want to estimate the effect of non-convex adjustment costs on investment behavior, the simplification is less justifiable as the last term in equations [9.6] and [9.7] can be of substantial size. This concern is of course supported by our knowledge of lumpiness in factor adjustment patterns.

Since it may not be reasonable to ignore the cost saving terms in equations [9.6] and [9.7] , we need to develop an approach that considers the effect that the level of investment chosen in period *t* has on the expectation of adjustments in future periods. In a paper published in 2004, Paola Rota deals with the possibility of corner solutions in the demand for labor, by considering a "discrete-time-discrete-choice dynamic structural model". The model estimates fixed costs of adjustment in a three-step approach where the first step estimates conditional adjustment probabilities that are required to obtain a structural parameter for a dynamic marginal productivity equilibrium condition similar to those in equations [9.6] and [9.7]. This condition takes into account that that the firm has a choice between adjustment and non-adjustment because of the existence of fixed costs in each time period *t*. After a structural parameter is obtained, it is used in estimating the fixed cost of labor adjustment.

The procedure is certainly less convenient than the standard approach where one simply ignores the cost saving issue that arises with non-convex costs, but shows that the issue can be dealt with empirically in a way that is consistent with the factor adjustment model in this thesis. In an empirical investigation of a model with interrelations, we would meet another challenge in treating the effect of interrelations when modeling discrete choice probabilities, but for now this problem is left out for a later treatment.

(A.2) Deriving Marginal Productivity Conditions

We start out by deriving an expression for λ_t^K and consider the same Lagrange function as from which we found optimal investment levels $\frac{I_t}{K_t}$ and $\frac{H_t}{L_t}$.

$$
\begin{aligned}
\mathbf{L} = {} & F(A_t, K_t, L_t) - w_t L_t - C(.) \\
& +\lambda_t^K \left(I_t + \left(1-\delta^K\right)K_t - K_{t+1}\right) \\
& +\lambda_t^L \left(H_t + \left(1-\delta^L\right)L_t - L_{t+1}\right) \\
& +\frac{1}{1+r}\cdot E_t \begin{bmatrix} F(A_{t+1}, K_{t+1}, L_{t+1}) - w_{t+1}L_{t+1} - C(.) \\ +\lambda_{t+1}^K\left(I_{t+1} + \left(1-\delta^K\right)K_{t+1} - K_{t+2}\right) \\ +\lambda_{t+1}^L\left(H_{t+1} + \left(1-\delta^L\right)L_{t+1} - L_{t+2}\right) \end{bmatrix} \\
& +\left(\frac{1}{1+r}\right)^2 \cdot E_t \begin{bmatrix} F(A_{t+2}, K_{t+2}, L_{t+2}) - w_{t+2}L_{t+2} - C(.) \\ +\lambda_{t+2}^K\left(I_{t+2} + \left(1-\delta^K\right)K_{t+2} - K_{t+3}\right) \\ +\lambda_{t+2}^L\left(H_{t+2} + \left(1-\delta^L\right)L_{t+2} - L_{t+3}\right) \end{bmatrix} + \left(\frac{1}{1+r}\right)^3 \cdot E_t[\quad]\ldots
\end{aligned}
\qquad [9.1]
$$

where

$$
\begin{aligned}
C(.) &= C\left(I_{t+s}, K_{t+s}, H_{t+s}, L_{t+s}\right) \\
&= \left(p_{t+s}^{I+} I_{t+s} \cdot \mathrm{I}\left(I_{t+s} > 0\right) + p_{t+s}^{I-} I_{t+s} \cdot \mathrm{I}\left(I_{t+s} < 0\right) + \alpha^K I_{t+s} + \frac{b^K}{2}\left(\frac{I_{t+s}}{K_{t+s}}\right)^2 K_{t+s}\right) \cdot \mathrm{I}\left(I_{t+s} \neq 0\right) \\
&\quad -\left(p_{t+s}^{H+} H_{t+s} \cdot \mathrm{I}\left(H_{t+s} > 0\right) + p_{t+s}^{H-} H_{t+s} \cdot \mathrm{I}\left(H_{t+s} < 0\right) + \alpha^L H_{t+s} + \frac{b^L}{2}\left(\frac{H_{t+s}}{L_{t+s}}\right)^2 L_{t+s}\right) \cdot \mathrm{I}\left(H_{t+s} \neq 0\right) \\
&\quad -\alpha^{KL} \cdot \mathrm{I}\left(H_{t+s} \neq 0\right) \cdot \mathrm{I}\left(I_{t+s} \neq 0\right)
\end{aligned}
$$

To obtain an expression for the shadow value λ_t^K, we set the derivative of L with respect to K_{t+1} equal to zero.

$$\frac{\partial \mathbf{L}}{\partial K_{t+1}} = 0$$

Taking the derivative gives us

$$-\lambda_t^K + E_t\left[\frac{1}{1+r}\left(\frac{\partial F\left(A_{t+1}, K_{t+1}, L_{t+1}\right)}{\partial K_{t+1}} + \frac{b^K}{2}\left(\frac{I_{t+1}}{K_{t+1}}\right)^2 + \lambda_{t+1}\left(1-\delta^K\right)\right)\right] = 0 \qquad [9.2]$$

which can be rewritten as

$$\lambda_t^K = E_t\left[\frac{1}{1+r}\left(\frac{\partial F\left(A_{t+1}, K_{t+1}, L_{t+1}\right)}{\partial K_{t+1}} + \frac{b^K}{2}\left(\frac{I_{t+1}}{K_{t+1}}\right)^2\right)\right] + E_t\left[\frac{\left(1-\delta^K\right)}{1+r}\lambda_{t+1}\right] \qquad [9.3]$$

Now, because λ_{t+s}^K can be represented as

$$\lambda_{t+s}^K = E_{t+s}\left[\frac{1}{1+r}\left(\frac{\partial F\left(A_{t+s+1}, K_{t+s+1}, L_{t+s+1}\right)}{\partial K_{t+s+1}} + \frac{b^K}{2}\left(\frac{I_{t+s+1}}{K_{t+s+1}}\right)^2\right)\right] + E_{t+s}\left[\frac{\left(1-\delta^K\right)}{1+r}\lambda_{t+s+1}\right]$$

[9.4]

we obtain

$$\begin{aligned}\lambda_t^K = {} & E_t\left[\frac{1}{1+r}\left(\frac{\partial F\left(A_{t+1}, K_{t+1}, L_{t+1}\right)}{\partial K_{t+1}} + \frac{b^K}{2}\left(\frac{I_{t+1}}{K_{t+1}}\right)^2\right)\right] \\ & + E_t\left[\frac{\left(1-\delta^K\right)}{\left(1+r\right)^2}\left(\frac{\partial F\left(A_{t+2}, K_{t+2}, L_{t+2}\right)}{\partial K_{t+2}} + \frac{b^K}{2}\left(\frac{I_{t+2}}{K_{t+2}}\right)^2\right)\right] \\ & + E_t\left[\frac{\left(1-\delta^K\right)^2}{\left(1+r\right)^3}\left(\frac{\partial F\left(A_{t+3}, K_{t+3}, L_{t+3}\right)}{\partial K_{t+3}} + \frac{b^K}{2}\left(\frac{I_{t+3}}{K_{t+3}}\right)^2\right)\right] + \ldots\end{aligned} \qquad [9.5]$$

and we get the following equality for the shadow value of capital:

$$\lambda_t^K = E_t\left[\sum_{s=0}^{\infty}\frac{\left(1-\delta^K\right)^s}{\left(1+r\right)^{s+1}}\left(\frac{\partial F\left(A_{t+s+1}, K_{t+s+1}, L_{t+s+1}\right)}{\partial K_{t+s+1}} + \frac{b^K}{2}\left(\frac{I_{t+s+1}}{K_{t+s+1}}\right)^2\right)\right] \qquad [9.6]$$

Similarly the shadow value of one unit of labor is:

$$\lambda_t^L = E_t\left[\sum_{s=0}^{\infty}\frac{\left(1-\delta^L\right)^s}{\left(1+r\right)^{s+1}}\left(\frac{\partial F\left(A_{t+s+1},K_{t+s+1},L_{t+s+1}\right)}{\partial L_{t+s+1}}+\frac{b^L}{2}\left(\frac{H_{t+s+1}}{L_{t+s+1}}\right)^2\right)\right] \qquad [9.7]$$

We assume that $Y_{t+s+1} = F\left(A_{t+s+1},K_{t+s+1},L_{t+s+1}\right) = A_{t+s+1}\left(L_{t+s+1}\right)^{\eta}\left(K_{t+s+1}\right)^{1-\eta}$, and that $P\left(W_{t+s}\right)$ is the probability of adjustment in period $t+s$. Leaving out the cost saving component from equation [eq.#], we have the following:

$$\begin{aligned}
&E_{t+s}\left[\frac{\partial F\left(A_{t+s+1},K_{t+s+1},L_{t+s+1}\right)}{\partial K_{t+s+1}}\right] = \left(1-\eta\right)E_{t+s}\left(\frac{Y_{t+s+1}}{K_{t+s+1}}\right)\\
&= P(W_{t+s})\frac{\partial F\left(\left(1-\delta^K\right)K_{t+s}+I_{t+s}\right)}{\partial K_{t+s+1}} + \left(1-P\left(W_{t+s}\right)\right)\frac{\partial F\left(\left(1-\delta^K\right)K_{t+s}\right)}{\partial K_{t+s+1}} \qquad [9.8]\\
&= P(W_{t+s})\left\{\frac{\partial F\left(\left(1-\delta^K\right)K_{t+s}+I_{t+s}\right)}{\partial K_{t+s+1}} - \frac{\partial F\left(\left(1-\delta^K\right)K_{t+s}\right)}{\partial K_{t+s+1}}\right\} + \frac{\partial F\left(\left(1-\delta^K\right)K_{t+s}\right)}{\partial K_{t+s+1}}
\end{aligned}$$

Further, we note that

$$\begin{aligned}
\frac{\partial F\left(A_{t+s+1},L_{t+s+1},\left(1-\delta^K\right)K_{t+s}\right)}{\partial K_{t+s+1}} &= \left(1-\delta^K\right)^{1-\eta}\frac{\partial F\left(A_{t+s+1},L_{t+s+1},K_{t+s}\right)}{\partial K_{t+s+1}}\\
&= \left(1-\eta\right)\left(1-\delta^K\right)^{1-\eta}\frac{Y_{t+s}}{\left(1-\delta^K\right)K_{t+s}} \qquad [9.9]\\
&= \left(1-\eta\right)\left(1-\delta^K\right)^{-\eta}\frac{Y_{t+s}}{K_{t+s}}
\end{aligned}$$

and use this to get

$$(1-\eta)E_{t+s}\left(\frac{Y_{t+s+1}}{K_{t+s+1}}\right)=(1-\eta)\left(1-\delta^{K}\right)^{-\eta}\frac{Y_{t+s}}{K_{t+s}}+P\left(W_{t+s}\right)\left\{\frac{\partial F\left(\left(1-\delta^{K}\right)K_{t+s}+I_{t+s}\right)}{\partial K_{t+s}}-\frac{\partial F\left(\left(1-\delta^{K}\right)K_{t+s}\right)}{\partial K_{t+s}}\right\}$$

, which in turn can be written as:

$$E_{t+s}\left(\frac{Y_{t+s+1}}{K_{t+s+1}}\right)=(1-\delta)^{-\eta}\cdot\frac{Y_{t+s}}{K_{t+s}}+\left(1-\delta^{K}\right)^{-\eta}P\left(W_{t+s}\right)\left\{\frac{Y_{t+s}}{K_{t+s}}\left[\left[\frac{\left(1-\delta^{K}\right)K_{t+s}}{\left(1-\delta^{K}\right)K_{t+s}+I_{t+s}}\right]^{\eta}-1\right]\right\}$$
$$=\left(1-\delta^{K}\right)^{-\eta}\left\{1+P\left(W_{t+s}\right)\left[\left[\frac{\left(1-\delta^{K}\right)K_{t+s}}{\left(1-\delta^{K}\right)K_{t+s}+I_{t+s}}\right]^{\eta}-1\right]\right\}\cdot\frac{Y_{t+s}}{K_{t+s}}$$

[9.10]

A corresponding derivation for the second term in the expression for λ_t^K can also be done by inserting discrete probabilities of investment, yielding;

$$E_{t+s}\left[\frac{b^K}{2}\left(\frac{I_{t+s+1}}{K_{t+s+1}}\right)^2\right]=P(W_{t+s})\frac{b^K}{2}\left(\frac{I_{t+s+1}}{K_{t+s+1}}\right)^2+\left(1-P\left(W_{t+s}\right)\right)0$$
$$=P(W_{t+s})\frac{b^K}{2}\left(\frac{I_{t+s+1}}{K_{t+s+1}}\right)^2$$

[9.11]

In the literature, because one has mainly focused on convex adjustment costs, one has assumed that $\left(\frac{I_t}{K_t}\right)^2$ is small. However, this assumption stands in contrast to our knowledge about micro-data behavior, which has been described as "lumps and bumps", and also the motivation behind this thesis which attempts to discuss a way of estimating adjustment costs that are likely to be non-convex.

First, let us see what the assumption of a small investment rate $\left(\frac{I_t}{K_t}\right)^2$ implicates, when we allow ourselves to write an approximation of the shadow value λ_t^K as

$$\lambda_t^K \approx E_t\left[\sum_{s=0}^{\infty}\frac{\left(1-\delta^K\right)^s}{\left(1+r\right)^{s+1}}\cdot\left(1-\eta\right)\left(1-\delta^K\right)^{-\eta}\left(\frac{Y_{t+s}}{K_{t+s}}\right)\right] \qquad [9.12]$$

$$\lambda_t^K \approx E_t\left[\sum_{s=0}^{\infty}\frac{\left(1-\eta\right)\left(1-\delta^K\right)^{s-\eta}}{\left(1+r\right)^{s+1}}\left(\frac{Y_{t+s}}{K_{t+s}}\right)\right] \qquad [9.13]$$

Using our previous result, we then obtain

$$\begin{aligned}
E_t\left(\frac{Y_{t+s}}{K_{t+s}}\right) &= \left(1-\delta^K\right)^{-\eta}\left\{1+P\left(W_{t+s-1}\right)\left[\left[\frac{\left(1-\delta^K\right)K_{t+s-1}}{\left(1-\delta^K\right)K_{t+s-1}+I_{t+s-1}}\right]^{\eta}-1\right]\right\}\times\frac{Y_{t+s-1}}{K_{t+s-1}} \\
&= \left(1-\delta^K\right)^{-\eta}\left\{1+P\left(W_{t+s-1}\right)\left[\left[\frac{\left(1-\delta^K\right)K_{t+s-1}}{\left(1-\delta^K\right)K_{t+s-1}+I_{t+s-1}}\right]^{\eta}-1\right]\right\}\times\ldots \\
&\ldots\times\left(1-\delta^K\right)^{-\eta}\left\{1+P\left(W_{t+s-2}\right)\left[\left[\frac{\left(1-\delta^K\right)K_{t+s-2}}{\left(1-\delta^K\right)K_{t+s-2}+I_{t+s-2}}\right]^{\eta}-1\right]\right\}\times\ldots \\
&\ldots\times\left(1-\delta^K\right)^{-\eta}\left\{1+P\left(W_{t+s-3}\right)\left[\left[\frac{\left(1-\delta^K\right)K_{t+s-3}}{\left(1-\delta^K\right)K_{t+s-3}+I_{t+s-3}}\right]^{\eta}-1\right]\right\}\times\ldots
\end{aligned}$$

$$\Rightarrow E_t\left(\frac{Y_{t+s+1}}{K_{t+s+1}}\right) = \left(1-\delta\right)^{-\eta(s+1)}\frac{Y_t}{K_t}\cdot\prod_{i=0}^{s}\left\{1+P\left(W_{t+i}\right)\left[\left[\frac{\left(1-\delta\right)K_{t+i}}{\left(1-\delta\right)K_{t+i}+I_{t+i}}\right]^{\eta}-1\right]\right\} \qquad [9.14]$$

If we further assume that $0<\beta<1$, and stick to the assumption that $\left(\frac{I_t}{K_t}\right)$ is small, then

$$\left\{1+P\left(W_{t+i}\right)\left[\left[\frac{(1-\delta)K_{t+i}}{(1-\delta)K_{t+i}+I_{t+i}}\right]^{\beta}-1\right]\right\}\approx 1$$

Which in turn results in the following approximation for λ_t^K, which is conveniently observable in many cases:

$$\lambda_t^K \approx \left[\sum_{s=0}^{\infty} \frac{(1-\beta)\left(1-\delta^K\right)^{s(1-\beta)-\beta}}{(1+r)^{s+1}} \left(\frac{Y_t}{K_t} \right) \right] \propto \theta \left(\frac{Y_t}{K_t} \right) \qquad [9.15]$$

While for labor, the expression becomes:

$$\lambda_t^L \approx \left[\sum_{s=0}^{\infty} \frac{(1-\beta)\left(1-\delta^L\right)^{s(1-\beta)-\beta}}{(1+r)^{s+1}} \left(\frac{Y_t}{L_t} \right) \right] \propto \theta \left(\frac{Y_t}{L_t} \right) \qquad [9.16]$$

(B) Descriptive Statistics of the Second Simulated Dataset (ii)

TABLE B.1: CORRELATION MATRIX

	H/L	I/K	L	K	q^K	q^L
H/L	1					
I/K	0.6558	1				
L	0.0060	0.0343	1			
K	-0.0563	-0.0291	0.5778	1		
q^K	0.1980	0.5538	0.0241	-0.0383	1	
q^L	0.5410	0.1615	-0.0329	-0.0633	0.5066	1

TABLE B.2: SUMMARY STATISTICS

Variable	Obs.	Mean	Std. dev.	Min.	Max.
H/L	1000	.011755	.1905979	-.7398592	.5436076
I/K	1000	.1934281	.2102485	-.5664857	.7501802
L	1000	358.0854	240.7971	25.46533	1389.317
K	1000	7884.791	3779.305	89.18464	20929.77
q^K	1000	3.084015	2.986225	-8.111652	12.74151
q^L	1000	3.916717	4.16364	-9.872766	20.57577

FIGURE B.1: HIRING

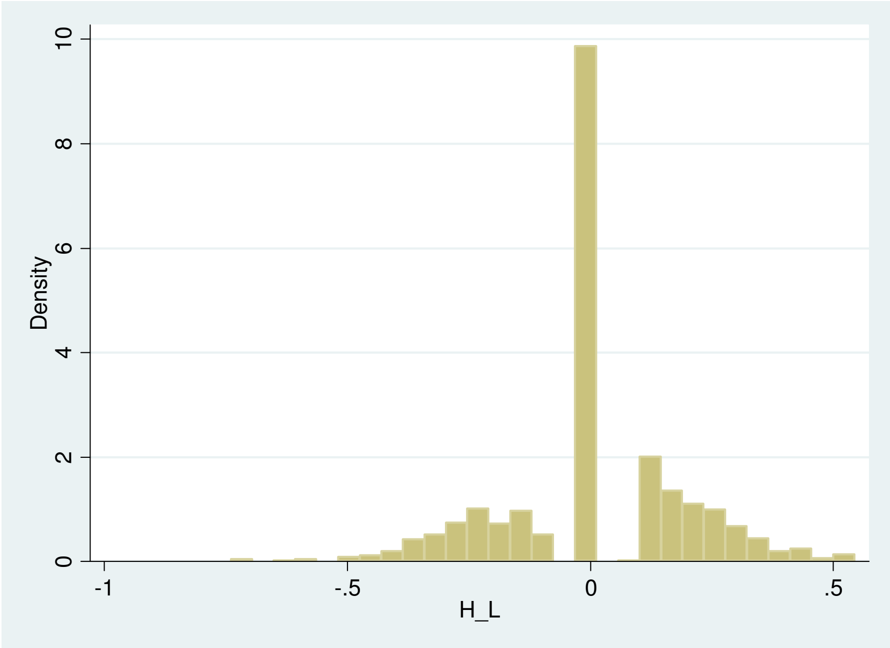


FIGURE B.2: INVESTMENTS

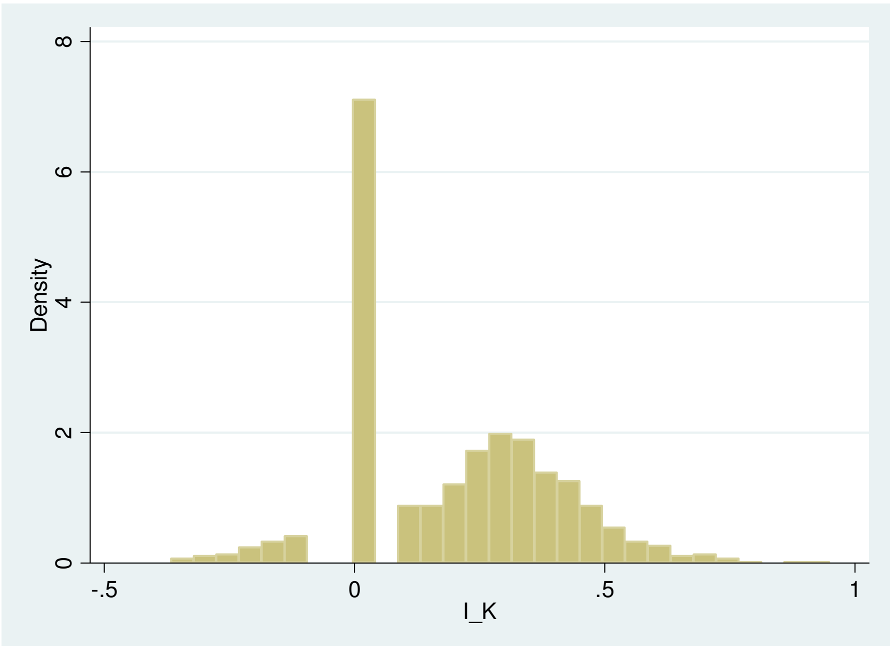


TABLE B.3: HIRING & INVESTMENT (JOINT DISTRIBUTION)

	Hiring			
Invest	-1	0	1	**Total**
-1	58	1	0	59
0	70	252	1	323
1	113	184	321	618
Total	241	437	322	1000