Optimal Investment Policy with Fixed Adjustment Costs and Complete Irreversibility

Nicolas Roys *
UCL, IFS, Paris 1

May 25, 2007

Abstract

This paper proves the optimality of an (S, s) policy in a discrete-time model of investment with fixed adjustment costs and complete irreversibility. Investment is shown to depend simultaneously on marginal and average Q, which are sufficient statistics of, respectively, marginal and total gains of adjustment. Cash-flows are not a correct proxy for the marginal value of capital, the latter being a non-monotonic function of profitability. Neither functional forms nor calibration are imposed and there is no need for numerical procedures. The result holds for a wide class of shocks and technologies. Proofs use the concept of K-convexity introduced by [Scarf, 1960].

JEL: C61, D21 and E2
Keywords: Investment, Adjustment Costs, Irreversibility, Dynamic Programming

1 Introduction

This paper proves the optimality of an (S, s) policy in a discrete-time model of investment with fixed adjustment costs and complete irreversibility under a minimal set of restrictions on technology and uncertainty. It is a well-known fact that investment is lumpy at the plant-level with long periods of inactivity punctuated by infrequent and large adjustment. For example in continuing and large U.S. manufacturing establishments, [Doms and Dunne, 1998] documents that more than half of them exhibit capital growth close to 50 percent in a single year and between 25 and 40 percent of an average plant's gross investment over the seventeen year period considered is concentrated in a single year period. To satisfyingly replicate observed microeconomic investment patterns, it is necessary to consider fixed adjustment costs and irreversibility [1]. The former are independent of the level of investment and are incurred whenever investment is nonzero. Thus a firm can avoid them by not investing at all. The latter depicts a resale price of capital inferior to its current acquisition price [2]. This paper considers that installed capital is valuable

*nicolas_r@ifs.org.uk. I would like to thank without implications Jean-Marc Robin, Cuong Le Van, François Gourio, Rabah Amir, Russell W. Cooper and Dan Hamermesh. Very useful feedback was also received from seminar and conference participants at the University of Paris 1, the University of Texas at Austin and the Society for Computational Economics.

[1] [Hamermesh and Pfann, 1996] and [Khan and Thomas, 2006]) survey the literature on adjustment costs. While still widely used for their tractability, convex adjustment costs models have fare poorly to account for observed investment behavior. The failure of the class of so-called partial adjustment models comes from the implied smooth and continual adjustment to shocks.

[2] because of *lemons effects*, capital specificity and market thiness.

only to the extent that it is used in production ; a situation called complete irreversibility. It describes many observed situations since plants sell capital very rarely. Instead, they reduce their capital stocks to lower desired levels by letting them depreciate. Scarce episodes of negative investment are usually interpreted as depreciation and obsolescence of existing capital stock. It may also concerns firms in distressed (see [Ramey and Shapiro, 2001] for a concrete example). In this case, capital sales and investments are two separate activities that do not interact. Only the second is part of regular productive activities. In most empirical analysis of investment, the dynamics of dis-investment is either estimates separately (see [Abel and Eberly, 2002] for example) or neglected.

The following definition gives a formal derivation of observed investment policy. I introduce the notation which will be more precisely defined in the rest of the paper.

Definition 1 *An investment policy, denoted i, is an (S, s) policy if there exists a target function and a threshold function, denoted S and s, that do not depend on past level of capital, denoted k, and such that*

$$i(A, k) = \begin{cases} 0 & \text{if } k > s(A) \\ S(A) - k & \text{if } k < s(A) \end{cases}$$

where A denote the state of the world

[Eberly and Van Mieghem, 1997] studies a special case where $S = s$ [3]. Such a situation arises when adjustment costs are linear but not fixed.

In an influential paper, [Abel and Eberly, 1994] develops a unified model of investment under uncertainty. Results are obtained once a precise stochastic process and an analytic profit-function are imposed. A drawback is the unavoidable normality assumption which renders the problem manageable. But in many economic situations, fat tails and asymmetries suggest that normality is too restrictive. And as stated by [Ackerberg et al., 2007]: *estimation of production functions has a long history in applied economics (...) Unfortunately, this history cannot be deemed an unqualifed success, as many of the econometric problems that hampered early estimation are still an issue today.* Characterizing optimal policies without relying on any functional forms is thus desirable. In comparison, the results hold here for a wide-class of technologies and with very few restrictions on the uncertainty confronting the firm. Also many of the subsequent papers use a geometric brownian motion as it is highly tractable. But the recent development of tests for units roots using panel data (see for example [Hall and Mairesse, 2003]) demonsrates that in most situations the presence of a unit root is rejected in favor of a process with very high persistence but inferior to unity. And the analysis is done in continuous-time. Yet, there is a growing literature that demonstrate the interests and advantages to work in discrete time. [Boyarchenko and Levendorskii, 2007] discusses that point: *discrete time is more natural for economics, and it allows one to obtain analytical results in some situations, where continuous-time models are either not applicable, or do not lead to simple analytical results.* For example, large investment episodes are often spread across two or three years ([Doms and Dunne, 1998]). While it is certainly involved to reproduce this in a continuous-time model, it emerges naturally by considering an appropriate time-interval in a discrete-time framework.

The empirical literature on investment dynamics is vast[4]. Since observations are available at equally spaced moments in time, it is naturally in discrete-time. In reduced-form

[3] This paper deals with positive and negative demand for inputs and an arbitrary number of factors.

[4] see [Bond and Reenen, 2007] for a recent survey.

empirical papers, it is customary to say that an (S, s) policy is optimal but the theoretical references are always continuous-time models. In structural models, several studies discuss the optimality of inaction and jumps in the investment process using numerical procedures[5]. But there is no formal proof: the (S, s) policy is rather observed in numerical examples. Here, no functional forms are imposed and there is no need for numerical procedures nor their associated calibration and simulations exercises. The characterization of an optimal policy with a state-dependent target and threshold can potentially help the identification strategy. Using indirect-inference, it helps understanding the crucial moments to match. Using maximum likelihood estimators, variants of the framework developed here were reduced to a dynamic discrete choice model to be amenable to a structural estimation[6]. A notable exception is [Pakes, 1994]. He showed how to handle both discrete and continuous control variables in a structural estimation. However, his methods involves considerable selection as it only considers adjustment periods in the estimation. With a full-characterization of the optimal policy, it is possible to structurally esimate the model using a generalized selection model where the value functions are approximated with Chebyshev Polynomials[7].

Despite the non-differentiability and the discontinuity of the one-period profit function, the value function is shown to be continuous and to be differentiable everywhere except at the threshold level where it admits unequals left and right derivatives. This is of interest since many authors have erroneously stated that the value function is non-differentiable. An analytical expression for the marginal value of capital is derived and interpreted. Two notable findings emerge: (i) the optimal investment policy depends simultaneously on the marginal value of capital and the total value of capital and (ii) the marginal value of capital is shown to be a non-monotonic function of profitability. These are strongly related to the litterature on investment - cash flows sensitivities. Most empirical work takes [Hayashi, 1982] neoclassical investment model as its point of departure. In this model, the firm is perfectly competitive, has a constant-returns-to-scale production function, and faces convex investment adjustment costs that are linearly homogeneous in investment and capital. [Hayashi, 1982]'s model implies that the optimal investment-capital ratio depends only on Tobin's average Q, defined as the ratio of the value of the firm to its capital stock. This prediction has generally been rejected by empirical studies, which show that cash flows and other measure of current profitability have a strong predictive power for investment after controlling for Tobin's Q. The common interpretation of these findings is the presence of financial frictions. Recent papers have challenged this interpretation. They compute dynamic general equilibrium models with financial frictions, calibrate them and look at the relation between Tobin's Q and investment in simulated series. [Gomes, 2001] show that financial frictions are neither necessary nor sufficient for significant cash-flow effects in standard regressions. [Cooper and Ejarque, 2001] and [Cooper and Ejarque, 2003] establish that decreasing returns and market power help to generate realistic correlations while financial frictions do not. [Bayraktar et al., 2005] show that the impact of external finance constraints on investment is relatively modest compared to the role of adjustment costs. Then much of the significance of cash flow variables in conventional estimates of Tobin's Q investment equations occurs because the strong assumptions necessary to make investment depend only on average Q do not hold in the data. If one follows this liter-

[5] the optimal decision is usually found after discretization of the state space by value function iteration.

[6] For example, [Rota, 2004] estimate a labor demand model with fixed costs using [Hotz and Miller, 1993]'s estimator. The former is suited for discrete decision processes (the decision to adjust or not). The continuous decision (how much to invest) is estimated non-parametrically.

[7] this is the approach of [Fuentes et al., 2006] in a model of investment with linear adjustment costs only.

ature, a solution (actually infeasible) would be to replace average Q by marginal Q in regressions. The present paper establishes, under very few assumptions, and whitout any financing constraints that it would still not be enough. Investment is shown to depends simultaneously on marginal and average Q, which are sufficient statistics of, respectively, marginal and total gains of adjustment. Further, the marginal value of capital is shown to be a non-monotonic function of profitability. This can explains why when one allows for non-linearity and/or uses measurement-error consistent estimators, the investment-cash flows correlations are significantly reduces.

Some results partly duplicate or generalize earlier related findings reported by [Caballero and Leahy, 1996 This very insightful paper solve a model with fixed costs and address many of the questions of the present paper. Yet, this is done with a very specific shock process and by imposing homogeneity assumptions. But many empirical studies show that assumptions about constant return to scale profits may not be correct. And they consider a scale-dependent fixed cost (proportional to capital) while this paper consider a true fixed costs. Recent empirical papers show that the latter are necessary to reproduce the different patterns of investment in large versus small firms (see below).

Even in times of growing computer power, the computational interest of this paper derserves attention. Under the assumptions made, when one solve numerically a model, he can restricts his attention to policies of the (S, s) form. For example, if one solves the model with value function iteration, it is sufficient to look at the first point of the capital sate space to find the optimal target (S). To find the threshold (s) it is sufficient to look at values below the target. Further, one can infer the value function at any point below the target from the computation of the value function at the target exactly.

This paper can also be of interest in two other fields. A lively litterature study how lumpy microeconomic investment decisions impact aggregated investment. A consensus has not yet been reached. The results found rely on simulations and are oftenly criticized regarding functional forms and calibration assumptions. This paper is a first step toward a less simulation-dependent analysis of this subject. Finally, a recent line of research relates stock return dynamics to firms' real investment decisions[8]. Using the results of this paper it would be possible to do so without assuming normality (which is strongly rejected when using commodity prices data).

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 derives the optimal decision rule in the case of complete irreversibility and provides a comparative statics analysis of the model. Section 4 establishes the differentiability of the value function and re-formulates investment policy in terms Tobin's average and marginal Q. Section 5 extends the results to non-stationary and multivariate shocks, a horizon of arbitrary length, time-varying parameters and flexible inputs. Section 6 concludes. All proofs are in the Appendix.

2 The Model

In this section, I introduce the model and impose assumptions under which the (S, s) policy is optimal but that are strong enough to be tractable for empirical work. In Section 5, I will relax some assumptions that have appeared too strong.

[8]For example, high book-to-market stocks benefits more than low book-to-market stocks from a forthcoming expansion in real activity. [Cooper, 2006] shows that this result is consistent with the presence of non-convex adjustment costs

2.1 Assumptions

Time is discrete and indexed by t. At each period, the firm decides to invest[9] or not ($i_t \geq 0$) over an infinite time horizon. Under the risk-neutral assumption, the firm's objective is to maximize its expected net present value which is the discounted sum of one-period profits net of adjustment costs . Her decision depends on A_t, an observable shock to the profitability[10] of the firm at the beginning of period t and k_t the level of capital inherited from the previous period. The one-period profit function is:

$$\pi(A,k,i) = R(A,k+i) - C(i) \tag{1}$$

$R(A,k)$ represents reduced-form profits and incorporates the optimal choice of flexible factors. It is based on an underlying production and costs functions where all flexible factors of production (labor, working time,...) have been optimized out (see Section 5.5.2 for a formal analysis). The simplest representation is: $R(A,k) = Ak^{\alpha} - rk$ where α is the curvature of the production function and r is the user-cost of capital. Usually, it is assumed that investment becomes productive with one period lag. Since it would complicate the notation without reversing the results[11], assume it is immediately productive.

$C(i)$ is the adjustment costs function:

$$C(i) = \begin{cases} F + pi & \text{if } i > 0 \\ 0 & \text{if } i = 0 \end{cases} \tag{2}$$

where F and p represent, respectively, fixed and linear adjustment costs. The fixed cost creates a discontinuity in $C(\cdot)$ at 0: $C(0) = 0$ while $\lim_{|k'-k|\to 0} C(k'-k) > 0$. To extend the analysis to general equilibrium and to consider heterogeneity in adjustment costs firms face, it is possible to consider the fixed cost as random accross firms so that the inaction range vary accross firms.

In a similar model, [Caballero and Leahy, 1996] consider a scale-dependent fixed cost. This is meant to reflect forgone profits due to the loss of production associated with installation of capital. Instead, this paper considers a true fixed-cost which means that it does not depend on the scale of operation. The main motivation is the stylised fact that smaller units are characterized by a substantially more intermittent investment. For example, [Nilsen and Schiantarelli, 2003] found that the frequency of zero investment for small units is more than two times than for large units[12]. And the evidence suggests that financial constraints are not the most likely explanation [13]. They conclude that the econometric evidence support the existence of a purely fixed cost component, unrelated to plant size[14].

I present the timing of the model and then the technical assumptions that will be used in the proofs.

[9] The results of the paper can be applied to a labor demand context. However, the assumed constancy of quits and adjustments costs (see below) would then be problematic since there is empirical evidence of their variations over the business cycle. That it why my focus is more on investment.

[10] Product demand, productivity of inputs and prices of flexibles factors (wages, energy prices,...) conditions are combined into this univariate shock. Multivariate extensions are considered in Section 5.1

[11] The consequences of a delay between the purchase and the availability of investment are well-understood. It increases the expected gap between actual and desired capital. When deciding whether to invest (or not) and the optimal amount of investment, the level of profitability considered is not the actual value but rather the expected value at the time where investment will be productive.

[12] This is also true for labor. [Nilsen et al., 2007] found that the frequency of episodes characterized by no employment changes decreases markedly with size.

[13] Precisely, they a found a similar pattern in small plant belonging to multi-plants large firms and also when they split the sample according to criteria correlated with the probability of financing constraints

[14] indivisibilities may also play a role

Assumption 1 *Timing of the model and State Space*

1. *At the beginning of the period, the manager knows his past input level ($k_t \in \mathcal{K} \subset \mathbb{R}^{+*}$) and current level of profitability ($A_t \in \mathcal{A}$) where $\mathcal{A}$ is a compact subset of $\mathbb{R}$*

2. *Given (k_t, A_t) the manager invests or not ($i_t \geq 0$)*

3. *The level of capital at the start of the next period, $t+1$, is:*

$$k_{t+1} = (1-\delta)(k_t + i_t) \tag{3}$$

   where δ is a positive and constant depreciation rate [15].

4. *New value of profitability is generated by a Markov process with transition function given by $Q : \boldsymbol{A} \times \mathcal{A} \longrightarrow [0,1]$ where $(\mathcal{A}, \boldsymbol{A})$ is a measurable space.*

Assumption 2 *The firm discounts future profits at a constant rate $\beta \in (0,1)$*

Assumption 3 *$Q(A_{t+1}, \cdot)$ is a continuous and strictly decreasing function for every $A_{t+1} \in \mathcal{A}$*

Assumption 4 *R is jointly continuous and concave for each $(A,k) \in (\mathcal{A}, \mathcal{K})$, $R(\cdot, k)$ is increasing for every $A \in \mathcal{A}$ and $R(A, \cdot)$ is single-peaked.*

A follows an exogenous and stationary markov process. An eventually assumed exogenous technological progress would render the problem nonstationary. However, [Bertsekas and Shreve, 1978] show how nonstationary models can be reduced to stationary ones by appropriate reformulation. Technology improvements and adoption can easily be introduced in the model. [Buettner, 2004] develops a parsimonious framework to analyse the distribution of future productivity conditional on R&D and current productivity. There is R&D investments by firms which influence the stochastic evolution of profitability at a cost that depend on the current level of profitability. Formally let the distribution of future profitability be $\tilde{Q}(A', \psi)$ and be stochastically increasing in R&D investments (denoted ψ). And, the cost of R&D investments is increasing in A. The results of the paper are unchanged to such a transformation. The assumption of exogeneity may appear as an important restriction on the learning process. Yet, it is consistent with the empirical literature[16]. An empirically relevant interpretation of Assumption 4 is that at the beginning of its life, the manager draw a permanent level of profitability from which actual profitability can deviates according to a markovian process. Precisely, Q can be thought as parametrized by this permanent level of profitability. Further, in Section 5.1, it is shown that the optimality of an (S, s) policy holds with almost no restrictions on the form of uncertainty.

From Assumption 3, $Q(A_{t+1}, \cdot)$ is stochastically increasing in the first-order stochastic dominance sense. There is a positive persistence of the shocks: high values of profitability today are more likely to be associated with high values of profitability tomorrow. It seems realistic and will allow comparative statics analysis. The transition function $Q(A', A)$ satisfies the Feller property. It guarantees that the expectation function, used later in

[15] for analytical convenience, depreciation is proportional to both accumulated capital and investment. This is consistent with the assumption that investment is immediately productive.

[16] [Pakes and Ericson, 1998] find that manufacturers sales appears to Markovian whereas retailers do not. As pointed by [Abbring and Campbell, 2005], Non-Markovian dynamics can not only arise from Bayesian learning but also from permanent and unobservable differences across entrepreneurs' choices of their firms' intended scales. Empirically, they find that heterogeneity across firm's pre-entry scale decisions and transitory shocks observed by entrepreneurs can fully explain surviving bars' long history dependence and no evidence of entrepreneurial learning.

the Bellman equation, is bounded and continuous. Note that the normality assumption is not necessary. It represents an improvement compared to continuous-time models where Brownian shocks and boundary conditions are necessary to obtain solutions.

The profit function includes flows payments on capital like for example maintenance costs or commitments associated to purchase of capital. Combined with decreasing returns and/or some degree of monopoly power, R is an increasing and then decreasing function of k. The more widely used functional forms is $Ak^\alpha - rk$ where α is the curvature of the revenue function and r is the cost of capital.

2.2 The Dynamic Programming Problem

Given the law of motion of capital, the sequence of investment $\{i_t\}_{t=0}^\infty$ is chosen to maximize the present discounted value of current and future profits:

$$\sup_{\{i_t\}_{t=0}^\infty} E\left[\sum_{t=0}^{\infty} \beta^t \pi(A_t, k_t, i_t)\right] \tag{4}$$

such that $k_{t+1} = (1-\delta)(k_t + i_t)$ and with k_0 given.

Before writing the model as a dynamic programming problem and its standard recursive formulation, it is necessary to deal with two technical issues: the objective function defined in (4) has to be well-defined and the capital state space to be compact. In the appendix, Proposition 8 guarantees the first and Proposition 9 deals with the second. Precisely, Proposition 9 shows that choices of k_t above a particular value (noted $\tilde{k}$) can not be optimal: an upper-bound to the capital state space exists. Since the capital stock can not take negative values, it has a natural zero lower bound. The compacity of the capital state (denoted $\mathcal{K}$) is derived from the optimal behavior of the firm and is not an ad-hoc assumption.

Define the value function at period t, $V(A_t, k_t)$ as the discounted expected value of current and future cash-flows:

$$V(A_t, k_t) = \sup_{\{i_j\}_{j=t}^\infty} E\left\{\sum_{j=t}^{\infty} \beta^{j-t} \pi(A_j, k_j, i_j) | k_t, A_t\right\} \tag{5}$$

From standard results in dynamic programming theory, I can focus on stationary and markovian policies and the value function $V(A, k)$ is given by the solution to Bellman's equation:

$$V(A, k) = \sup_{0 \le i \le \tilde{k} - k} \{W(A, k+i) - C(i)\} \tag{6}$$

where

$$W(A, k) \equiv R(A, k) + \beta \int_{\mathcal{A}} V(A', k(1-\delta)) Q(\mathrm{d}A', A) \tag{7}$$

In a sense, W can be interpreted as the value of the firm when in the considered period investment is not possible and the optimal policy is followed in all future periods. While next Proposition may appear technical, it is of interest because it shows that the presence of the fixed costs does not affect the existence and unicity of a solution Bellman's equation and more importantly the value function is continuous despite the discontinuity in the one-period profit function. Note that Proposition 1 requires little assumption on the stochastic process.

Proposition 1 *V is unique, jointly continuous in (A, k) and increasing in A*

Proof see Appendix A $\square$

Following a standard argument, in the remaining of the paper, the operator sup is replaced by the operator max. Let $O : \mathcal{A} \times \mathcal{K} \to \mathcal{K}$ denote the policy correspondence that corresponds to the set of solution to (6):

$$O(A,k) = \Big\{k' \in [k, \tilde{k}] | V(A,k') = W(A,k') - C(k'-k)\Big\} \tag{8}$$

Corollary 1 *$O(A,k)$ is non-empty, compact-valued and upper-hemi-continuous*

3 Optimal Decision Rule

3.1 Mathematical Background

The discontinuity of the one-period profit function (due to the fixed adjustment cost) implies that the value function is not concave. Traditional dynamic programming arguments are not available. Nevertheless the concept of K-concavity introduced by [Scarf, 1960] can be successfully used[17]. There is two equivalent ways to define K-concavity.

Definition 2

1. *a real-valued function g is called F-concave for $F \geq 0$, if for any $z \geq 0$, $b > 0$ and any y,*
$$g(y+z) - F \leq g(y) + \frac{z}{b}(g(y) - g(y-b))$$

2. *a real-valued function g is called F-concave for $F \geq 0$, if for any $x_0 \leq x_1$ and $\lambda \in [0,1]$,*
$$g((1-\lambda)x_0 + \lambda x_1) \geq (1-\lambda)g(x_0) + \lambda g(x_1) - \lambda F$$

Next Proposition proof can be found in [Bertsekas and Shreve, 1978], except the last property whose proof is in apprendix B.

Proposition 2

1. *A concave function is 0-concave and hence also F-concave for all $F \geq 0$*

2. *If $\{g_n(x)\}$ is a sequence of F-concave functions and $g = \lim_{n\to\infty} g_n$ is the pointwise limit of these functions, and if $|g(x)| < \infty$ for all x, then g is F-concave*

3. *If $f(x)$ is F-concave in $x \in [0, \bar{x}]$, where $F \geq 0$, then the function*
$$g(x) = \max_{y \geq x} \{f(y) - FI\{y > x\}\}$$
is also F-concave

[17] in the remaining of the paper, I use the term F-concavity instead of K-concavity to avoid any confusion with k.

3.2 Optimality of an (S, s) Policy

Essentially, I tailor the arguments for a class of commodity price speculation problems in [Hall and Rust, 2005] to derive the optimality of an (S,s) policy in the investment problem.

Proposition 3 *For any $(A, k) \in \mathcal{A} \times \mathcal{K}$, the functions V and W are F-concave in k and the optimal decision rule $i(A, k)$ takes the form of a state-dependent (S, s) policy:*

$$i(A,k) = \begin{cases} 0 & \text{if } k > s(A) \\ S(A) - k & \text{if } k < s(A) \end{cases} \tag{9}$$

where the functions $S(A)$ and $s(A)$ are given by:

$$S(A) = \min \arg \max_{0 \leq k \leq \tilde{k}} [W(A,k) - pk] \tag{10}$$

$$s(A) = \min \Big(k \in [0, \tilde{k}] | W(A,k) \geq W(A, S(A)) - p[S(A) - k] - F \Big) \tag{11}$$

The value function V can be expressed as:

$$V(A,k) = \begin{cases} W(A, S(A)) - F - p[S(A) - k] & \text{if } k \in [0, s(A)] \\ W(A,k) & \text{otherwise} \end{cases} \tag{12}$$

Proof see Apprendix C □

Recall that the optimal policy depicts in Proposition 3 coincides with the observed patterns of investment: several periods of inaction, when $k < s(A)$, followed by bursts of capacity adjustment, when k jumps to $S(A)$.

Defining a target $S(A)$ and a threshold $s(A)$ is noteworthy. With this in mind, the optimization problem is separated in two parts: choosing an optimal level of investment ignoring the fixed cost[18], and deciding whether or not to incur the fixed costs and invest at all. Note that $s(A) \leq S(A)$ because $s(A)$ is defined as the smallest point where optimal investment is zero. An important contribution of this paper is that no functional forms (nor calibration) are imposed and there is no need for numerical procedures to obtain the result.

To deal with multiple maximas, the target $S(A)$ is defined as the smallest value of the policy correspondence $O(A)$. In Section 4, it turns out that the target is unique. But differentiability of the value function will be proven first.

3.3 Interpretation and Graphical Illustration

For illustrative purposes, I plot the value of inaction (noted W or V^i) as a concave function of k (the green line). The value of ordering (noted V^a) is a linear function of k (the blue line). The value function is the maximum of the two curves. The linearity is due to linear adjustment costs. At the target level, the gap between the value of inaction and the value of an (optimal) adjustment is exactly equals to the fixed adjustment cost. If there is only fixed adjustment costs, the value of ordering would equate the value of inaction evaluated at the target level.

If there is an adjustment, the level of capital is chosen independently of its lagged value. In other words, there is no partial adjustment: if the manager decides to adjust, he directly jumps to the target without smoothing. To understand this property,

[18]it does not appear in the definition of $S(A)$. Yet it impacts on $S(A)$ through its effects on the value function.

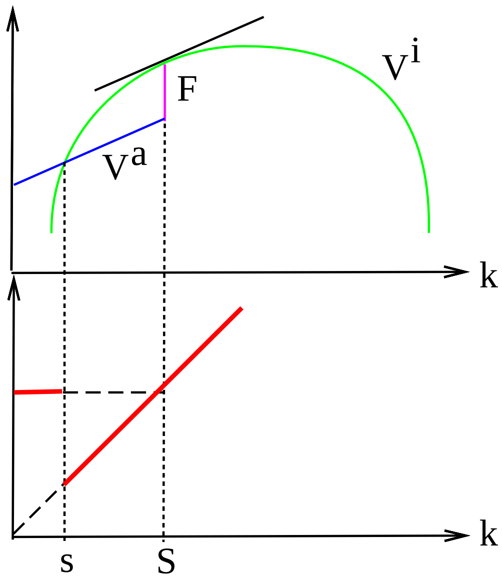

Figure 1: Value Function and Optimal Decision Rule

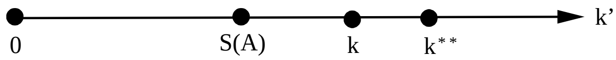

Figure 2: Value Function and Optimal Decision Rule

consider the definition of the investment policy when there is an adjustment: $i(A,k) = \min \arg\max_{0<i\leq \tilde{k}-k}[W(A,k+i) - pi]$. Using a simple variable change $k' = k+i$, k does not appear in the objective function $\max_{k<k'\leq \tilde{k}}[W(A,k') - pk']$ but it does in the set over which to optimize: $k' \in]k, \widetilde{k}]$. Indeed, this is a problem only if $S(A) < k$. A situation like the one in the following picture is conceivable: if a negative investment level were allowed, it would be the arg max of the objective function). But as it is not, the firm prefers to invest (and go to the maximum over the state space considered, say k^{**}) rather than staying inactive.

However this situation is not possible. For all $(A,k,i) \in \mathcal{A} \times (S(A), \tilde{k}] \times (0, \tilde{k} - k]$, $W(A,k) - pk$ being F-concave, it holds:

$$W(A,k+i) - p(k+i) - F \quad = \quad W(A,k) - pk + \frac{z}{k - S(A,k)}[W(A,k) - pk - W(A,S(A)) + pS(A)]$$

Since the term on the right hand side of the above inequality is non-positive, the value of inaction $W(A,k)$ is bigger than the value of investing any positive amount.

Between $s(A)$ and $S(A)$, the marginal gain of adjusting employment is superior to its marginal costs (p). Graphically, the slope of the green curve is bigger than the slope of the blue curve [19]. However, inaction is optimal because the total gain of adjusting employment is inferior to its total cost. After $S(A)$, both total and marginal gains are inferior to their respective costs.

[19] The differentiability of the value function is established in Section 4

At the level of previous capital in which the firm is indifferent between adjustment and non-adjustment, the level of capital chosen, if the firm adjusts, is superior to previous capital. Linear adjustment costs (or complete irreversibility) alone do not produce this phenomena. This is really due to the presence of a fixed cost the firm aims to avoid incurring too frequently. Further the magnitude of adjustment (when it occurs) must be large enough to create a total profit gain sufficient to offset the total cost. When there is no fixed cost, the decision-marker evaluates marginal gains only.

3.4 Comparative Statics

This section looks at the reply of the optimal invesment and the probability of investing to an increase in profitability using the general theory developed by [Topkis, 1978]. The following supermodularity assumption says that the marginal productivity of capital is increasing in profitability.

Assumption 5 *R is supermodular for each $(A, k) \in (\mathcal{A}, \mathcal{K})$*

Proposition 4 *S and s are non-decreasing functions.*

Proof see Appendix D □

Proposition 4 illustrates a simple principle: when profitability is high, the optimal level of investment and the probability of investing is higher. The intution is that the marginal value of capital increases in productivity and that a firm with higher current productivity is more likely to have better realisations of productivity in the future. Following a positive shock, the firm becomes more willing to invest (the threshold s rises) and wants to invest more (the target S rises).

A full-analysis of the effects of exit and outside options on the investment behavior would go beyond the scope of the present paper. A simple way to incorporate exit is to consider that firms face a probability of receiving a shock to the profitability level (say π) which would cause an exit of the market or the realloction of capital to any other outside activities. The gains from such a situation can be any function of the capital stock. Then the effects of a rise in this probability or of the gains of such a situation can be analysed as the impact of a decrease in the discount rate β.

Proposition 5 *If $\beta^1 < \beta^2$, $S(A; \beta^1) \leq S(A; \beta^2)$ and $s(A; \beta^1) \leq s(A; \beta^2)$ for each $A \in \mathcal{A}$.*

Proof see Appendix D □

Using Proposition 5, a firm which anticipate a rise in its probability of leaving the market reduces its optimal level of investment and its probability of investing.

4 Marginal and Average Value of Capital

The aim of this section is to characterize the optimal policy in terms of the marginal and average value of capital, also known as marginal and average Q. As an intermediary step, the value function is shown to be differentiable *a.e.* except at the threshold level where it admits unequals left and right derivatives. This is an interesting result per se since it holds despite the non-differentiability and the discontinuity of the one-period profit function. Further many authors have erroneously stated that the value function is non-differentiable.

It is obviously necessary to assume differentiability of the one-period profit function to expect differentiability of the value function.

Assumption 6 *R is differentiable and strictly concave*

Proposition 6 *The value function V is continuously differentiable in k almost everywhere on the domain $\mathcal{K}\backslash s(A)$ for every $A \in \mathcal{A}$. The partial derivative obeys the equation:*

$$V_k(A,k) = R_k(A, k+i(A,k)) + \beta(1-\delta)\int V_k(A', (1-\delta)(k+i(A,k)))Q(\mathrm{d}A', A) \quad (13)$$

which can be decomposed as:

$$V_k(A,k) = \begin{cases} p & \text{if } k < s(A) \\ R_k(A,k) + \beta(1-\delta)\int V_k(A',(1-\delta)k)Q(\mathrm{d}A',A) & \text{if } k > s(A) \end{cases} \quad (14)$$

Corollary 2 *For every $A \in \mathcal{A}$, the target $S(A)$ is unique and satisfies*

$$R_k(A, S(A)) + \beta(1-\delta)\int V_k(A', S(A)(1-\delta))Q(\mathrm{d}A', A) = p \quad (15)$$

Corollary 3 *The functions S and s are unique, continuous and increasing*

Proof see Appendix E □

The interpretation of Corollary 2 is the usual one. If the firm decides to invest, it equates marginal gains from adjustment to its marginal costs. Note that the fixed cost does not appear explictly in the preceding equation. However, it impacts the target indirectly through its effects on the value function. Corollary 3 establishes the unicity of the target for each level of profitability. Thus the optimal decision rule is a policy and not a correspondence. This is an improvement compared to [Caballero and Leahy, 1996] and [Caballero and Engel, 1999] where the strict concavity of the one-period profit function do not imply unicity.

To compute the derivative with respect to capital of the value function, I solve recursively the function W. Since for every $A \in \mathcal{A}$, $|W(A,0)| < \infty$, when j goes to infinity:

$$\lim_{j\to\infty} \beta^j \int W(A_j, k(1-\delta)^j) \prod_{l=1}^{j-1} Q\left(\mathrm{d}A_l, A_{l-1} | k(1-\delta)^l \geq s(A_l)\right) = 0$$

and since $s(\cdot)$ is invertible (using Corollary 3), it holds:

$$\begin{aligned} W(A_0,k) &= R(A_0,k) \\ &+ \sum_{j=1}^{\infty} \beta^j \left(\prod_{l=1}^{j-1} Q([\underline{A}, s^{-1}(k(1-\delta)^l)], A_{l-1})\right) \Big(\int_{\underline{A}}^{s^{-1}(k(1-\delta)^j)} R(A_j, k(1-\delta)^j) Q(\mathrm{d}A_j, A_{j-1}) \\ &+ \int_{s^{-1}(k(1-\delta)^j)}^{\bar{A}} \left[W(A_j, S(A_j)) - pS(A_j) + pk(1-\delta)^j - F\right]\Big) Q(\mathrm{d}A_j, A_{j-1}) \end{aligned} \quad (16)$$

The marginal value of capital can be written:

$$
\begin{aligned}
& W_k(A_0,k) \\
= \;& R_k(A_0,k)+\sum_{j=1}^{\infty}(\beta(1-\delta)^j)(\prod_{l=1}^{j-1}Q([\underline{A},s^{-1}(k(1-\delta)^l)],A_{l-1})\cdot \\
& [\int_{\underline{A}}^{s^{-1}(k(1-\delta)^j)}R_k(A_j,k(1-\delta)^j)+\int_{s^{-1}(k(1-\delta)^j)}^{\bar{A}}p]Q(\mathrm{d}A_j,A_{j-1}) \\
+\;& s_k^{-1}(k(1-\delta)^j)q(s^{-1}(k(1-\delta)^j),A_{j-1})\cdot[R(s^{-1}(k(1-\delta)^j),k(1-\delta)^j) \\
& -[W(s^{-1}(k(1-\delta)^j),S(s^{-1}(k(1-\delta)^j)))-pS(s^{-1}(k(1-\delta)^j))+pk(1-\delta)^j-F]] \\
+\;& \prod_{l=2}^{j-1}\frac{q(s^{-1}(k(1-\delta)^l),A_{l-1})}{Q([\underline{A},s^{-1}(k(1-\delta)^l)],A_{l-1})}\cdot(\int_{\underline{A}}^{s^{-1}(k(1-\delta)^j)}R(A_j,k(1-\delta)^j)Q(\mathrm{d}A_j,A_{j-1}) \\
& +\int_{s^{-1}(k(1-\delta)^j)}^{\bar{A}}[W(A_j,S(A_j))-pS(A_j)+pk(1-\delta)^j-F]Q(\mathrm{d}A_j,A_{j-1}))
\end{aligned}
$$

Using the preceding equation, the marginal value of capital can be separated in three components. The first component is the discounted sum of marginal product of capital times the probability of non-adjustment. The effect of an additional unit of capital lasts when the firm adjusts and chooses a new level of capital, the target, which has been proven independent of past values of capital. So one additional unit of capital has an effect until the point in time where there is another adjustment. At this exact point in time, one additional unit of capital reduces by p the amount of linear adjustment costs. The preceding effect would be the only one if probabilities of investing in the future were constant. However, changes in k alter the time of the next adjustment. The two other effects takes this into account. An increase in capital decreases the probability of an adjustment in the future. There is a lost of the future marginal gains to adjust. Yet by definition, if the firm actually adjusts both total gains of adjustment and the marginal gains of adjustment are superiors to their respective costs and are positive. So if there is an adjustment, the first effect dominates the second one. The third effect is positive. By increasing capital, the firm reduces its probabilities of adjusting in the future which increase the number of period where the accumulated capital will have an impact on the marginal value of the firm.

The optimal policy can be formulated in terms of the marginal and the average value of capital. Define the average value of capital, also know as Tobin's Q: $Q(A,k)=\frac{W(A,k)}{pk}$. Define the marginal value of capital, also known as marginal q, as $q(A,k)=\frac{W_k(A,k)}{p}$. Using the preceding results, it holds:

Proposition 7 *For every* $(A,k)\in\mathcal{A}\times\mathcal{K}$

$$
i(A,k)=\begin{cases} 0 & \text{if } Q(A,k)>1+Q(A,S(A))-\frac{S(A)}{k}-\frac{F}{pk} \\ q^{-1}(A,1) & \text{otherwise} \end{cases} \tag{17}
$$

The interpretation is the following. The probability of an adjustment depends only on Tobin's average Q. The latter is a sufficient statistics for total gains of an adjustment. It entirely explains the decision to adjust or not. Note that the fixed cost is divided by the level of the capital stock which captures the level of operation. It would then follows that fixed cost matters more for small firms. The model repoduces the stylised facts that frequency of inaction decreases with size (cf. supra). The optimal amount of investment is entirely determined by Tobin's marginal q, which is a sufficient statistics for the marginal gains of an adjustment.

An important question is the link between marginal q and cash-flows. Given the level of capital at the beginning of the period, these two statistics are correlated with profitability. It is fair enough to say that cash-flows are positively correlated with profitability for a given value of capital at the beginning of the period. In the presence of fixed cost, the link between marginal q and profitability is ambiguous. While this paper proves it formally and in a very general framework, the intuition comes from [Caballero and Leahy, 1996]: *a positive profitability shock raises the marginal profitability of capital and the incentive to invest. It is the future marginal profits that do not necessarily rise. Because investment is lumpy, future marginal profits fall when investment is imminent, so that q falls as investment approaches.* The preceding discussion can be formally stated by looking at the effect of an increase in profitability on the marginal value of capital. The marginal value of capital can be re-written:

$$\begin{aligned}
W_k(A,k) &= R_k(A,k) \\
&+ \beta \int_{\underline{A}}^{S^{-1}(k(1-\delta))} W_k(A,k) Q(\mathrm{d}A_j, A_{j-1}) \\
&+ \beta \int_{S^{-1}(k(1-\delta))}^{s^{-1}(k(1-\delta))} W_k(A,k) Q(\mathrm{d}A_j, A_{j-1}) \\
&+ \beta \int_{s^{-1}(k(1-\delta))}^{\bar{A}} p Q(\mathrm{d}A_j, A_{j-1}) \qquad (18)
\end{aligned}$$

A positive profitability shock increases reduced-form profits. This is the standard intuitive effect. The three other effects are related to the impact on future profitability distribution. Given that higher values of profitability are more likely in the near future, the probability of an adjustment in the near future increases. However, if the firm adjusts, the marginal value of capital is equated to p. And given A, for every $k \in [s(A), S(A)), W(A,k) > p$: the marginal value of capital is higher than p. So because of the fixed costs, it may happen that the marginal value of capital is higher in a non-adjustment situation compared to a situation where the firm adjusts. An increase in profitability today put more weights to profitability levels in which the firm adjust and less weights to profitability levels in which the marginal value of capital is higher than p. Then, the marginal value of capital may decrease with the level of profitability in some portions of the state space. Precisely this situation is more likely when the firm is already close to an adjustment i.e. k is higher than $s(A)$ but very close to. Consequently, the link between the marginal value of capital and profitability is non-monotonic. This translates into a non-monotonic relationship between the marginal value of capital and cash-flows.

This theoretical point gives a possible explanation for the non-robust correlation between cash flows and - investment. High cash-flows are a proxy for high values of profitability. If high cash-flows correspond to a situation of high profitability and low accumulated capital, then investment will effectively be high if the firm cross the threshold. But high-cash flows may also reflect a situation where both profitability and accumulated are at their average level. This idea is that when capital at the beginning of the period is close to the threshold level $s(A)$, the link between cash-flows and investment may be negative.

5 Extensions

5.1 A More General Framework

The degree of generality in the main text is chosen to be amenable to a structural estimation of the model. I show how to transpose all the results of this paper in a more general framework which allows non-stationary and multivariate shocks, a horizon of arbitrary length and time-varying parameters. Consider the probability space $(\Omega, \mathcal{F}, P)$ and filtration $\mathbb{F} = \bigcup_t \mathcal{F}_t$ where $\mathcal{F}_t$ represents the information available at time t. Given the choice of factor $k_t \in \boldsymbol{R}^+$ and the state of the world $w \in \Omega$ [20] The one-period profit function in period t is denoted $R_t(\omega, k_t)$. For each $t \in \{1, \ldots, T\}$, the function $R_t(\cdot, \cdot)$ is jointly measurable, $R_t(\cdot, k_t)$ is $\mathcal{F}_t$-measurable for each $k_t \in \boldsymbol{R}^+$ and $R_t(\omega, \cdot)$ is concave for each $\omega \in \Omega$. The adjustment costs function can be generalized to a function with time-varying parameters: $C(i_t) = p_t i_t + F_t I\{i_t \neq 0\}$. It is however necessary to assume that the discounted fixed cost sequences $\{\beta^{t-1} F_t\}_{t=1}^{T+1}$ are non-increasing in t [21].

Under standard regularity conditions, for all $t \in \{1, \ldots, T\}$, the value function $V_t(\cdot, \cdot)$ is $\mathcal{F}_t$-measurable and is given by the solution to Bellman's equation:

$$V_t(w_t, k_t) = \sup_{i_t \geq 0} \{R_t(w_t, k_t + i_t) - C_t(i_t) + \beta E\left[V_{t+1}\left(w_{t+1}, (k_t + i_t)(1 - \delta)\right) |\mathcal{F}_t\right]\}$$

The final-period value function $V_{T+1}(\cdot, \cdot)$ is assumed jointly measurable and $V_{T+1}(A, \cdot)$ is F_{T+1}-concave for each $A \in \mathcal{A}$. To avoid end-of-horizon investment effects, it is possible to assume: $V_{T+1}(\cdot, \cdot) = 0$. It is straightforward to see that all the results of the paper still holds in the more general framework presented here.

Finally, it may seem restrictive to require that $R(A, \cdot)$ is concave in that it is imposed for every level of profitability. [Amir, 1997] presents a generalization of stochastic technology where the production process exhibits decreasing returns to scale in some average sense but it may enjoy increasing returns for some favorable resolutions of the uncertainty. To the price of further notations and regularity assumptions, the results of the present paper can be extended to such a situation.

5.2 Flexible Inputs

For the optimal investment policy analysis, flexible (in the sense of adjusted at no costs) inputs are unimportant because their choice is purely static: it has no dynamic implications. Since every period the firm chooses its optimal level at no cost, past level of these inputs will not impact on their current level. However, It is necessary to assume that after having made the optimal choice of these factors, the one-period profit function remains concave in (A, k) jointly. For example, add to the preceding model a single flexible factor (say l) which enter the one-period profit function and is costless to adjust. Also to be more realistic[22], assume that some part of the shock is not observed until after the investment decision is made and that the flexible factor is chosen after the rigid one. Formally, consider a point in time $t + b$ between period t and $t + 1$ ie. $b \in (0, 1)$. Assume, that during t and $t + b$, the firms gets new informations on the profitability of the current period: an *i.i.d* shock that affect profits (say μ) occurs. Denote by $\tilde{R}(A, k, l, \mu)$ the considered profit-function and assumed it is jointly continuous and concave in its arguments. The first-order condition for l is $R_l = 0$. Using the implicit function theorem, $l_k^*(A, k, \mu) = -\frac{\tilde{R}_{lk}(A,k,l,\mu)}{\tilde{R}_{ll}(A,k,l,\mu)}$. When

[20] So far, I have assumed a univariate stochastic process $A = \{A_t(\omega) : \omega \in \Omega, t \geq 0\}$.

[21] It holds trivially when these costs are fixed over-time.

[22] Notably, it avoids a deterministic relatonship between the two factors.

the profit function has positive (negative) crosspartials so that its inputs are economic complements (substitutes), a higher optimal level of the dynamic inputs justifies a higher (lower) level of the static input. This can explain the negative correlation of hours and employment growth at the plant level observed in [Cooper et al., 2004]. Applied to this framework, the number of employees is the rigid factor and working time is the flexible one. Following a positive shock, if the number of employees stays constant (because of adjustment costs), then working time increases to accomodate the shocks. But if the firm hires new workers then workng-time eventually stays at its regular-level.

5.3 Multi-factors

The generalization to a multi-factors situation requieres important restrictions on the profit and adjustment costs functions. It is necessary to generalize the concept of K-concavity to a n-dimensional Euclidean space. [Gallego and Sethi, 2005] propose the following definition:

Definition 3 *A function $g : \mathbb{R}^n \to \mathbb{R}$ is $(F_0, F_1, \ldots, F_n)$-concave if*

$$g\left((1-\lambda)x_0 + \lambda x_1\right) \geq (1-\lambda)g(x_0) + \lambda g(x_1) - \lambda F(x_1 - x_0)$$

for all x_0, x_1 with $x_0 \leq x_1$ and all $\lambda \in [0,1]$ and where the function $\boldsymbol{F} : \mathbb{R}^n_+ \to \mathbb{R}_+$ is defined as follows:

$$\boldsymbol{F}(x) = F_0 I\{x > 0\} + \sum_{l=1}^{n} F_l I\{x_l > 0\}$$

The key question is the following. Does $(F_0, F_1, \ldots, F_n)$-concavity persists after a dynamic programming iteration ? Consider the model of the present paper where $i, k \in \mathbb{R}^n_{+*}$ and n is the number of inputs. The adjustment costs function becomes: $\boldsymbol{C}(i) = F_0 I\{i > 0\} + \sum_{l=1}^{n} F_i I\{i_l > 0\} + pi$ where p is the vector of inputs prices. The other assumptions and notations are unchanged. Using the results of [Gallego and Sethi, 2005], the answer is positive in two specific cases:

1. $\boldsymbol{F}(x) = F_0 I\{x > 0\}$
2. $\boldsymbol{F}(x) = \sum_{l=1}^{n} F_l I\{x_l > 0\}$ and $R(A,k) = \sum_{l=1}^{n} R_l(A, k_l)$ where $R_l(A, k_l)$ satisfies Assumption 4 for every $l \in 1, \ldots, n$

In the first case, the same fixed cost is payed whenever an input is changed: $F_0 \geq 0$ and $F_i = 0$, $i = 1, \ldots, n$. So there is no input-specific fixed-cost. In the second-case, there is no joint-fixed cost: $F_0 = 0$ and the demands are independent. Both are very strong restrictions and more results are expected in the future. Yet, the results of the paper hold straightforwardly in the two depicted situations.

It is then possible to transpose the analysis of [Eberly and Van Mieghem, 1997] and [Dixit, 1997]. The interaction between dynamic factors is similar to the interaction between flexible and dynamic factors (see section 5.2). An adjustment in one dynamic input increases the marginal productivity of all the others inputs. Then the likelihood of adjusting every other inputs increases. The following figure illustrates the determination of the optimal decision rule in the case of two dynamic inputs (say k_1 and k_2). This is a generalization of [Eberly and Van Mieghem, 1997] to a case with positive fixed costs but a drawback of the analysis here is that negative demand for inputs are not allowed. Figure 3 shows the space is divided in 4 regions. When the level of the two inputs is low, the

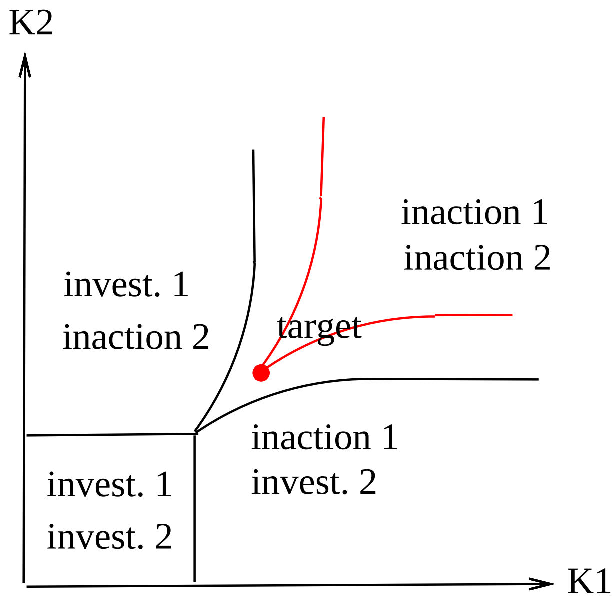

Figure 3: Optimal decision rule with $n = 2$

optimal policy is to invest in both inputs up to reach the red point. In the region in the upper-right, inaction is optimal with respect to both inputs because the reward of changing the input quantities is insufficient. The figure shows that a firm will be all the more willing to invest in one input when it holds a relatively important quantity of the other input.

6 Conclusion

The main contribution of the paper must be found as a stepping-stone for future work. By proving the high degree of generality of (S, s) policies, it provides a parsimonious framework to study and estimate various problem of investment and capital accumulation. Unicity, continuity and differentiability (except at the threshold) of the value function is established. The optimal policy is unique and fully characterized. Investment is shown to depends simultaneously on marginal and average Q, which are sufficient statistics of, respectively, marginal and total gains of adjustment. The marginal value of capital is shown to be a non-monotonic function of profitability. The former forbide the use of cash-flows coefficients as a test of financing constraints and re-estalish in a simple and relevant model that marginal Q is different from average Q.

References

[Abbring and Campbell, 2005] Abbring, J. H. and Campbell, J. R. (2005). A firm's first year. Tinbergen Institute Discussion Papers 05-046/3, Tinbergen Institute.

[Abel and Eberly, 1994] Abel, A. B. and Eberly, J. C. (1994). A unified model of investment under uncertainty. *American Economic Review*, 84(5):1369–84.

[Abel and Eberly, 2002] Abel, A. B. and Eberly, J. C. (2002). Investment, q with fixed costs: An empirical analysis. mimeo.

[Ackerberg et al., 2007] Ackerberg, D., Benkard, L., Berry, S., and Pakes, A. (2007). *Econometric Tools for Analyzing Market Outcomes*, volume 6, chapter Handbook of Econometrics.

[Amir, 1997] Amir, R. (1997). A new look at optimal growth under uncertainty. *Journal of Economic Dynamics and Control*, 22(1):67–86.

[Bayraktar et al., 2005] Bayraktar, N., Sakellaris, P., and Vermeulen, P. (2005). Real versus financial frictions to capital investment. Working Paper Series 566, European Central Bank.

[Bertsekas and Shreve, 1978] Bertsekas, D. P. and Shreve, S. E. (1978). *Stochastic Optimal Control: The Discrete Time Case*.

[Bond and Reenen, 2007] Bond, S. and Reenen, J. V. (2007). *Microeconometric Models of Investment and Employment*, volume 6 of *Handbook of Econometrics*.

[Boyarchenko and Levendorskii, 2007] Boyarchenko, S. and Levendorskii, S. (2007). Practical guide to real options in discrete time. *International Economic Review*, 48(1).

[Buettner, 2004] Buettner, T. (2004). Rd and the dynamics of productivity. mimeo, London School of Economics.

[Caballero and Engel, 1999] Caballero, R. J. and Engel, E. M. R. A. (1999). Explaining investment dynamics in u.s. manufacturing: A generalized (s,s) approach. *Econometrica*, 67(4):783–826.

[Caballero and Leahy, 1996] Caballero, R. J. and Leahy, J. V. (1996). Fixed costs: The demise of marginal q. NBER Working Paper 5508, National Bureau of Economic Research, Inc.

[Cooper, 2006] Cooper, I. (2006). Asset pricing implications of nonconvex adjustment costs and irreversibility of investment. *Journal of Finance*, 61(1):139–170.

[Cooper and Ejarque, 2001] Cooper, R. and Ejarque, J. (2001). Exhuming q: Market power vs. capital market imperfections. NBER Working Papers 8182, National Bureau of Economic Research, Inc.

[Cooper and Ejarque, 2003] Cooper, R. and Ejarque, J. (2003). Financial frictions and investment: Requiem in q. *Review of Economic Dynamics*, 6(4):710–728.

[Cooper et al., 2004] Cooper, R. W., Haltiwanger, J. C., and Willis, J. (2004). Dynamics of labor demand: Evidence from plant-level observations and aggregate implications. NBER Working Papers 10297, National Bureau of Economic Research, Inc.

[Dixit, 1997] Dixit, A. (1997). Investment and employment dynamics in the short run and the long run. *Oxford Economic Papers*, 49(1):1–20.

[Doms and Dunne, 1998] Doms, M. E. and Dunne, T. (1998). Capital adjustment patterns in manufacturing plants. *Review of Economic Dynamics*, 1(2):409–429.

[Eberly and Van Mieghem, 1997] Eberly, J. C. and Van Mieghem, J. A. (1997). Multi-factor dynamic investment under uncertainty. *Journal of Economic Theory*, 75(2):345–387.

[Fuentes et al., 2006] Fuentes, O., Gilchrist, S., and Rysman, M. (2006). Discrete adjustment costs and investment dynamics: a maximum likelihood approach. mimeo, Boston University.

[Gallego and Sethi, 2005] Gallego, G. and Sethi, S. P. (2005). K-convexity in rn. *Journal of Optimization Theory and Applications*, 127(1):71–88.

[Gomes, 2001] Gomes, J. F. (2001). Financing investment. *American Economic Review*, 91(5):1263–1285.

[Hall and Mairesse, 2003] Hall, B. H. and Mairesse, J. (2003). Testing for unit roots in panel data: An exploration using real and simulated data. mimeo, UC Berkeley.

[Hall and Rust, 2005] Hall, G. and Rust, J. (2005). The (s,s) policy is an optimal trading strategy in a class of commodity price speculation problems. *Economic Theory*, forthcoming.

[Hamermesh and Pfann, 1996] Hamermesh, D. S. and Pfann, G. A. (1996). Adjustment costs in factor demand. *Journal of Economic Literature*, 34(3):1264–1292.

[Hayashi, 1982] Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica*, 50(1):213–24.

[Hotz and Miller, 1993] Hotz, J. V. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic model. *Review of Economic Studies*, 69:497–529.

[Khan and Thomas, 2006] Khan, A. and Thomas, J. K. (2006). *Adjustment Costs*. The New Palgrave Dictionary of Economics. Palgrave Macmillan.

[Nilsen and Schiantarelli, 2003] Nilsen, O. and Schiantarelli, F. (2003). Zeros and lumps in investment: Empirical evidence on irreversibilities and nonconvexities. *The Review of Economics and Statistics*, 85(4):1021–1037.

[Nilsen et al., 2007] Nilsen, O. A., Salvanes, K. G., and Schiantarelli, F. (2007). Employment changes, the structure of adjustment costs, and plant size. *European Economic Review*, 51(3):577–598.

[Pakes, 1994] Pakes, A. (1994). *Dynamic structural models, problems and prospects: mixed continuous discrete controls and market interaction*, volume 2 of *Advances in Econometrics, Sixth World Congress*. C A Sims, cambridge university press edition.

[Pakes and Ericson, 1998] Pakes, A. and Ericson, R. (1998). Empirical implications of alternative models of firm dynamics. *Journal of Economic Theory*, 79(1):1–45.

[Ramey and Shapiro, 2001] Ramey, V. A. and Shapiro, M. D. (2001). Displaced capital: A study of aerospace plant closings. *Journal of Political Economy*, 109(5):958–992.

[Rota, 2004] Rota, P. (2004). Estimating labor demand with fixed costs. *International Economic Review*, 45(1):25–48.

[Scarf, 1960] Scarf, H. E. (1960). *The optimality of (S,s) policies in the dynamic inventory problem*. Mathematical Methods in the Social Sciences. Stanford University Press.

[Stokey et al., 1989] Stokey, N. L., Lucas, R. J. E., and Prescott, E. C. (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.

[Topkis, 1978] Topkis, D. M. (1978). Minimizing a submodular function on a lattice. *Operations Research*, 26(2):305–321.

A Proof of Proposition 1

Proposition 8 *(4) exists and is less than infinity*

Proof Consider a plant which faces no adjustment costs. Then, It easy to see that solving (4) is equivalent to solve a sequence of static decision problems. $R(A, \cdot)$ has a global maximizer for every $A \in \mathcal{A}$ (by the Weierstrass Theorem) and the expected discounted profits of such a plant is finite (R being bounded). Since it represents an upper-bound of the supremum, if it exists, the latter must be finite. To show that the supremum exists, it is sufficient to show that there exists one investment sequence which yields a value greater than $-\infty$. The sequence $i_t = 0$ for all $t \geq 0$ yields a return of 0. □

Given Proposition 8, I can define the supremum function associated to (4), say v^*.

Proposition 9 *For every $A \in \mathcal{A}$, k never exceed $\tilde{k}(A)$ defined as $R(A, \tilde{k}) - F - p\tilde{k} + E\left[\sum_{t=1}^{\infty} \beta^t R(A_t, k_t + i_t)\right] = 0$.*

Proof Let $\{k_t(A_t)\}_{t=0}^{\infty}$ be a sequence such that $\Pr\left[k_t(A_t) > \tilde{k}(A_t)\right] > 0$ for some $t > 0$. Then I prove that $v^* > \sum_{j=0}^{\infty} \beta^j \Pi(A_j, k_j, i_j)$. Without loss of generality, suppose that $k_1 = i_0 > \tilde{k}(A_1)$. From the definition of $\tilde{k}(A_1)$, $R(A_1, k_1) - F - pk_1 + E\left[\sum_{t=1}^{\infty} \beta^t R(A_t, k_t + i_t)\right] < 0$. Because $\sup_{\{i_t\}_{t=1}^{\infty}} E\left[\sum_{t=1}^{\infty} \beta^t \Pi(A_t, k_t, i_t)\right] \leq E\left[\sum_{t=1}^{\infty} \beta^t R(A_t, k_t + i_t)\right]$ and $v^* \geq 0$ (see Proof of Proposition 8), it follows:

$$v^* > R(A_1, k_1) - F - pk_1 + \sup_{\{i_t\}_{t=1}^{\infty}} E\left[\sum_{t=1}^{\infty} \beta^t \Pi(A_t, k_t, i_t)\right]$$

□

The following set are defined under the usual sup-norm $\|\cdot\|$. Let $B_{\mathcal{AK}}$ be the set of bounded function $V : \mathcal{A} \times \mathcal{K} \longrightarrow \boldsymbol{R}$. Let $C_{\mathcal{AK}}$ be the set of continuous function $V : \mathcal{A} \times \mathcal{K} \longrightarrow \boldsymbol{R}$. Denote T the operator associated to Bellman equation 6: $T(V) = V$. The value function can be re-written as: $V = \max\left\{V^a, V^i\right\}$ where $V^a(A, k) = \sup_{0 \leq i \leq \tilde{k} - k} \{W(A, k + i) - C(i)\}$ and $V^i(A, k) = W(A, k)$. I show that the operator T maps $C_{\mathcal{AK}}$ into itself a unique fixed point V in $C_{\mathcal{AK}}$. It is easy to show that T satisfies Blackwell's sufficient conditions for a contraction mapping. Since $B_{\mathcal{AK}}$ is complete, T has a unique fixed point V in $B_{\mathcal{AK}}$. Consider $V \in C_{\mathcal{AK}}$. By Lemme 9.5 in [Stokey et al., 1989] and Assumption 3, $V \in C_{\mathcal{AK}} \Rightarrow \left[\int_{\mathcal{A}} V(A', (1 - \delta)(k + i)) Q(\mathrm{d}A', A)\right] \in C_{\mathcal{AK}} \Rightarrow V^i \in C_{\mathcal{AK}}$. By the theorem of the maximum, $V^a(A, k) \in C_{\mathcal{AK}}$. The operator maxbeing continuous, $T(V) \in C_{\mathcal{AK}}$ and T maps $C_{\mathcal{AK}}$ into itself. Under Assumptions 3 and 4, $V(\cdot, k)$ is increasing for every $A \in \mathcal{A}$.

B Proof of Proposition 7.3

Define $m(x_i) = \inf \arg\max_{y \geq x_i} (f(y)), i \in \{0,1\}$ and $x_\lambda = (1-\lambda)x_0 + \lambda x_1$.

The proof uses repeateadly the following inequalities:

$$\begin{aligned} m(x_i) &\geq x_i && \text{for all } i \in (0,1) \\ g(x) &\geq f(x) && \text{for all } x \in [0, \bar{x}] \\ g(x) &> f(y) - k && \text{for all } y > x \end{aligned}$$

It is necessary to consider four different cases.

(1) $m(x_i) = x_i, i \in \{0,1\}$. In this case, $g(x_i) = f(x_i), i \in \{0,1\}$. $f(x)$ being F-concave, it holds:

$$\begin{aligned} g(x_\lambda) &\geq f(x_\lambda) \\ &\geq (1-\lambda)f(x_0) + \lambda f(x_1) - \lambda F \end{aligned}$$

(2) $m(x_i) > x_i, i \in \{0,1\}$. In this case, $(1-\lambda)m(x_0) + \lambda m(x_1) \geq x_\lambda$, $g(x_i) = f(m(x_i)) - F, i \in \{0,1\}$ and $m(x_0) \leq m(x_1)$. With this in mind and using the F-concavity of f, it holds:

$$\begin{aligned} g(x_\lambda) &\geq f((1-\lambda)m(x_0) + \lambda m(x_1)) - F \\ &\geq (1-\lambda)f(m(x_0)) + \lambda f(m(x_1)) - F - \lambda F \\ &\geq (1-\lambda)(f(m(x_0)) - F) + \lambda(f(m(x_1)) - F) - \lambda F \end{aligned}$$

(3) $m(x_0) = x_0$ and $m(x_1) > x_1$. In this case, $g(x_0) = f(x_0)$ and $g(x_1) = f(m(x_1)) - F$. There exists $\mu \leq \lambda$ s.t. $x_\lambda = (1-\mu)x_0 + \mu m(x_1)$. Then,

$$\begin{aligned} g(x_\lambda) &\geq f(x_\lambda) \\ &\geq (1-\mu)f(x_0) + \mu f(m(x_1)) - \mu F \\ &= (1-\lambda)g(x_0) + \lambda(g(m(x_1) + \lambda F - \lambda F + (\lambda - \mu)(g(x_0) - f(m(x_1)) + F) \end{aligned}$$

(4) $m(x_1) = x_1$ and $m(x_0) > x_0$. In this case, $g(x_1) = f(x_1)$ and $g(x_0) = f(m(x_0)) - F$

(4.1) $x_\lambda \leq m(x_0)$

$$\begin{aligned} g(x_\lambda) &\geq f(m(x_0)) - F \\ &\geq (1-\lambda)f(m(x_0)) - F) + \lambda f((m(x_0)) - \lambda F \\ &\geq (1-\lambda)g(x_0) + \lambda f((m(x_0)) - \lambda F \end{aligned}$$

The proof is completed by noting that $f((m(x_0)) = \max_{y > x_0} f(y) \geq f(x_1) = g(x_1)$

(4.2) $x_\lambda > m(x_0)$. There exists $\mu \leq \lambda$ s.t. $x_\lambda = (1-\mu)x_0 + \mu m(x_1)$. Then,

$$\begin{aligned} g(x_\lambda) &\geq f(x_\lambda) \\ &\geq (1-\mu)f(m(x_0)) + \mu f(x_1) - \mu F \\ &= (1-\lambda)g(x_0) + \lambda g(x_1) + \lambda F - \lambda F + (\lambda - \mu)(g(x_0) - f(m(x_1)) + F) \end{aligned}$$

C Proof of Proposition 3

Lemmas 1-4 constitute Proposition 3's Proof. Define the operators Γ and Λ by:

$$\begin{aligned} \Gamma(W)(A,k) &= \max_{0\le i\le \tilde{k}-k} [W(A,k+i) - C(i)] \\ \Lambda(V)(A,k) &= R(A,k) + \beta \int_{\mathcal{A}} V(A',k) f(\mathrm{d}A'|A) \end{aligned} \tag{19}$$

Let $\mathcal{F}_{\mathcal{AK}}$ denote the class of functions $V(A,k)$ which are continuous [23] in A and k, and F-concave as a function of $k \in \mathcal{K}$ for all $A \in \mathcal{A}$. Lemma 1 proves that if $W(A,k)$ is F-concave, so is V. Lemma 2 proves the optimality of an (S,s) policy. Lemmas 3 and 4 show that the composition operators $\Lambda \circ \Gamma$ and $\Gamma \circ \Lambda$ map $\mathcal{F}_{\mathcal{AK}}$ into $\mathcal{F}_{\mathcal{AK}}$. The proof of Lemma 2 directly follows from Lemma 3 in [Hall and Rust, 2005] and is therefore omitted.

Lemma 1 $\Gamma : \mathcal{F}_{\mathcal{AK}} \to \mathcal{F}_{\mathcal{AK}}$. *That is, if W(A,k)* $\in \mathcal{F}_{\mathcal{AK}}$, *then V(A,k)* $\in \mathcal{F}_{\mathcal{AK}}$.

Proof Consider W(A,k) $\in \mathcal{F}_{\mathcal{AK}}$. It implies $W(A,k) - pk \in \mathcal{F}_{\mathcal{AK}}$. Note that

$$V(A,k) = (\max_{k'} \left[W(A,k') - pk' - FI\{k' > k\}\right]\Big) + pk$$

where the first term is F-concave as a function of $k \in \mathcal{K}$ for all $A \in \mathcal{A}$ by Lemma **??** and so is the second term. $\square$

Lemma 2 *Suppose the function W(A,k) is continuous and F-concave in k for all A. Then the solution to the functional equation (6) has the following decision rule:*

$$i(A,k) = \begin{cases} 0 & \text{if } k > s(A) \\ S(A) - k & \text{if } k < s(A) \end{cases} \tag{20}$$

where the functions $S(A)$ and $s(A)$ are given by:

$$S(A) = \min \arg \max_{0 \le k \le \tilde{k}} [W(A,k) - pk] \tag{21}$$

$$s(A) = \min \left\{ k \in [0,\tilde{k}] | W(A,k) \ge W(A,S(A)) - p[S(A) - k] - F \right\} \tag{22}$$

Lemma 3 *Under the preceding assumptions,* $\Lambda \circ \Gamma : \mathcal{F}_{KC} \to \mathcal{F}_{KC}$

Proof By Lemma 1 if $U \in \mathcal{F}_{KC}$, then $V = \Gamma(U) \in \mathcal{F}_{KC}$. By Lemma **??**, there exist functions S and $s : \mathcal{A} \to R$ satisfying $0 \le s(A) \le S(A) \le \tilde{k}$ for which $V = \Gamma(U)$ can be represented as

$$\Gamma(U)(A,k) = \begin{cases} U(A,S(A)) - F - p(S(A) - k) & if k \in [0, s(A)] \\ U(A,k) & otherwise \end{cases}$$

Since positive linear combinations and point-wise limits of F-concave functions are F-concave, it follows that $\int_{\mathcal{A}} V(A', k(1-\delta)) f(\mathrm{d}A'|A)$ is F-concave in k over $\mathcal{K}$. Given that $\Lambda \circ \Gamma(U)(A,k) = AR(k) + \beta \int_{\mathcal{A}} V(A', k(1-\delta) f(\mathrm{d}A'|A)$, $\Lambda \circ \Gamma(U)(A,k)$ is F-concave in $k \in \mathcal{K}$ for all $A \in \mathcal{A}$. $\square$

[23] Note that continuity of V and W has already been proved.

Lemma 4 *Under the preceding assumptions, the function V and W are F-concave function of $k \in \mathcal{K}$ for all $A \in \mathcal{A}$.*

Proof

Since $\Gamma \circ \Lambda(V)$ is a contraction mapping, the fixed point $V = \Gamma \circ \Lambda(V)$ can be uniformly approximated by the method of successive approximations starting from an initial guess, $V_0 = 0$. $\Lambda(V_0) = AR(k)$ is concave in k by assumption 4. So, $\Lambda(V_0) \in \mathcal{F}_{KC}$. Lemma 1 implies that $V_1 = \Gamma \circ \Lambda(V_0) \in \mathcal{F}_{KC}$. Lemma 3 implies that $\Lambda(V_1) = (\Lambda \circ \Gamma) \circ \Lambda(V_0) \in \mathcal{F}_{KC}$. Continuing inductively, we see that for each $t \geq 0$ in the sequence of successive approximations, $V_t \in \mathcal{F}_{KC}$. Since the fixed point V is a uniform limit of functions in $\mathcal{F}_{KC}$ it follows that $V \in \mathcal{F}_{KC}$. Since $W = \Lambda \circ \Gamma(V)$, Lemma 3 also implies that $W \in \mathcal{F}_{KC}$. $\square$

D Proof of Proposition 4 and Proposition 5

The proof establishes that the value function is supermodular in (A, k) and is supermodular in (k, β). The results then follows from Theorem 6.1 in [Topkis, 1978].

Lemma 5 *$\tilde{C}(k', k) = C(k' - k)$ is submodular in (k', k) for every $k' \geq k$*

Proof the proof directly follows from the linearity of $C(k' - k)$ for every $k_1 < k_2 \leq k_1' < k_2'$ $\square$

Lemma 6 *$V(A, k)$ is supermodular in (A, k)*

Proof Let $S_{\mathcal{AK}} \subset C_{\mathcal{AK}}$ be the set of continuous, supermodular function $V : \mathcal{A} \times \mathcal{K} \longrightarrow \boldsymbol{R}$ and such that $V(\cdot, k)$ is non-decreasing for every $A \in \mathcal{A}$ and each $k \in \mathcal{K}$. I show that T maps $S_{\mathcal{AK}}$ into itself. Consider $V \in S_{\mathcal{AK}}$. V can be written as $V(A, k') = \max_{k' \in [k, \tilde{k}]} W(A, k') - C(k' - k)$. For a function on $\mathbb{R}^2$, increasing differences is equivalent to supermodularity. I need to show that $\int V(A', k')F(\mathrm{d}A', A)$ satisfies increasing differences in (A, k'). Consider $A_1 > A_2$ and $k_2' > k_1'$. It holds: $V(A_2, k_2') - V(A_2, k_1') \geq V(A_1, k_2') - V(A_1, k_1')$ which means that the function $V(\cdot, k_2') - V(\cdot, k_1')$ is non-decreasing for every $A \in \mathcal{A}$ and $k_2' > k_1'$. Combined with $F(A', \cdot)$ is stochastically increasing, it follows:

$$\begin{aligned}
\int \left[V(A', k_2') - V(A', k_1')\right] Q(\mathrm{d}A', A_2) &\geq \int \left[V(A', k_2') - V(A', k_1')\right] Q(\mathrm{d}A', A_1) \\
\int V(A', k_2')Q(\mathrm{d}A', A_2) - \int V(A', k_1')Q(\mathrm{d}A', A_2) &\geq \int V(A', k_2')Q(\mathrm{d}A', A_1) - \int V(A', k_1')Q(\mathrm{d}A', A_1)
\end{aligned}$$

which means that $\int V(A', k')Q(\mathrm{d}A', A)$ satisfies increasing differences in (A, k'). Combined with $R(A, k')$ supermodular in (A, k') (by assumption) and$-C(k' - k)$ is supermodular in (k, k') (by Lemma 5), $W(A, k') - C(k' - k)$ in (A, k', k). Maximizing over k' preserves supermodularity in the remaining variables (A, k) ([Topkis, 1978]). Consequently, $V = T(V)(A, k)$ is supermodular in (A, k). $\square$

Following a similar approach it can be shown that:

Lemma 7 *$W(A, k) - pk$ is supermodular in (A, k)*

Lemma 8 *$V(A, k; \beta)$ is supermodular in (β, k) for each $A \in \mathcal{A}$*

Proof Using the definition of the value function, it is obvious that $V(A,k;\beta^1) \le V(A,k;\beta^2)$. Following the logic of Lemma 6, consider $V(A,k;\beta)$ supermodular in (k,β). Then, for each $A \in \mathcal{A}$ and for every $k_1 < k_2$,

$$V(A,k_2;\beta^2) - V(A,k_1;\beta^2) \quad \ge \quad V(A,k_2;\beta^1) - V(A,k_1;\beta^1)$$

Multiplying the left-hand-side by β^2 and the right-hand-side by β^1 preserves the inequality and shows that $\beta V(A,k_1;\beta)$ is supermodular in (k,β). R and C being independent of β, the result follows by induction. $\square$

From [Topkis, 1978], the set of optimal solutions is non-decreasing in A and in β. $S(A)$ being the lowest element of the set, Proposition 4 and Proposition 5 follow.

Consider $A' \in \mathcal{A}$, there exists $\tilde{A} \in \mathcal{A}$ such that $S(\tilde{A}) = s(A')$. It follows that for every $A \in [0,\tilde{A}]$, $s(A) \le S(A) \le S(\tilde{A}) \le s(A')$.

E Proof of Proposition 12

Consider a sequence of $j+1$ period problems, $j = 0,1\ldots,J \in \mathbb{N}$ with value functions satisfying: $V^{j+1}(A,k) = \max_{\{k_t\}_{t=0}^{j+1}} E_0 \sum_{t=0}^{j} \beta^t \pi(A_t,k_t,k_{t+1}-k_t)$. The sequence $V^{j+1}(A,k)$ is generated by iterating on $V^0(A,k)$ with the operator T, i.e.,

$$V^{j+1}(A,k) = \max_{k^j \ge k} R(A,k^j) - C(k^j - k) + \beta \int V^j(A',k^j)Q(\mathrm{d}A',A) = T^{j+1}V^0 \qquad (23)$$

where k^j is the optimal choice of capital for the $j+1$ period problem. Clearly $V^0(A,k) = 0$. For each $j = 0,1,\ldots,J$ define $W^j(A,k) = R(A,k) + \beta \int V^j(A',k)Q(\mathrm{d}A',A)$ and accordingly $S^j(A,k)$ and $s^j(A)$. Using Proposition 3, it holds:

$$\begin{aligned} V^{j+1}(A,k) &= \begin{cases} W^j(A,S^j(A)) - F - pS^j(A) + pk & \text{if } k \le s^j(A) \\ W^j(A,k) & \text{if } k \ge s^j(A) \end{cases} \\ k^j(A,k) &= \begin{cases} S^j(A) & \text{if } k < s^j(A) \\ k & \text{if } k > s^j(A) \end{cases} \end{aligned}$$

Lemma 9 *For $j \ge 0$, $V^{j+1}(A,k)$ is continuously differentiable with respect to k almost everywhere on the domain $\mathcal{K} \backslash s^j(A)$ for every $A \in \mathcal{A}$ with a partial derivative $V_k^{j+1}(A,k)$ satisfying:*

$$V_k^{j+1}(A,k) = R_k(A,k^j) + \beta(1-\delta) \int V_k^j(A',k^j)Q(\mathrm{d}A',A) \qquad (24)$$

and can be decomposed as:

$$V_k^{j+1}(A,k) = \begin{cases} p & if k < s^j(A) \\ R_k(A,k) + \beta(1-\delta)\int V_k^j(A',(1-\delta)k)Q(\mathrm{d}A',A) & if k > s^j(A) \end{cases} \qquad (25)$$

Proof I shall show that $V^{j+1}(A,k)$ is continuously differentiable in k for each fixed A provided that $V^j(A,k)$ is continuously differentiable in k for each fixed A.

If $k^j(A) > k$, the first-order necessary condition for the maximum problem on the right-hand side is:

$$R_k(A,k^j) - p + \beta(1-\delta) \int V_k^j(A',(1-\delta)k^j)Q(\mathrm{d}A',A) = 0$$

and [?]'s theorem implies that $V^{j+1}(A,k)$ is differentiable in k with derivative given by: $V_k^{j+1}(A',k) = p$. Combined with the first-order condition, it gives Equation 9.

If $k^j(A) = k$, $V^{j+1}(A,k) = R(A,k) + \beta \int V^j(A',k)F(\mathrm{d}A',A)$ and consequently, $V^{j+1}(A,k)$ is differentiable in k with derivative given by Equation 24. □

Corollary 4 *For all $j \geq 0$, $S^j(\cdot)$ and $s^j(\cdot)$ are increasing*

Proof For $j \geq 0$, for every $A \in \mathcal{A}$, the target S^j(A) satifies:

$$R_k(A, S^j(A)) + \beta(1-\delta) \int V_k^j(A', (1-\delta)S^j(A))Q(\mathrm{d}A', A) = p$$

Consider the case $j = 0$. Since $R_{kA} > 0 and R_{kk} < 0$, $S^{0\prime}(A) = -\frac{R_{kA}(A,S^0(A))}{R_{kk}(A,S^0(A))} > 0$ and $s^{0\prime}(A) = -\frac{R_A(A,s^0(A)) - R_A(A,S^0(A))}{R_k(A,s^0(A)) - p} > 0$. Now consider $j \geq 0$. Provided that $V_{kk}^j(A,k) \leq 0$, using Proposition 9 and $R_{kk}(A,k) < 0$,

$$V_{kk}^{j+1}(A,k) = R_{kk}(A,k^j) + \beta \int V_{kk}^j(A',k^j)F(\mathrm{d}A',A) < 0$$

Differentiating the first-order condition, it holds:

$$S^{j\prime}(A) = -\frac{W_{kA}^j(A, S^j(A))}{W_{kk}^j(A, S^j(A))} > 0$$

Differentiating $W^j(A, s^j(A)) - W^j(A, S^j(A)) + F + pS^j(A) - ps^j(A) = 0$, it holds:

$$s^{j\prime}(A) = -\frac{W_A^j(A, s^j(A)) - W_A^j(A, S^j(A))}{W_k^j(A, s^j(A)) - p} > 0$$

□

Proposition 10 *The value function V is continuously differentiable in k almost everywhere on the domain $\mathcal{K}$*
$\{s(A)\}$ for every $A \in \mathcal{A}$. The partial derivative obeys the equation:

$$V_k(A,k) = R_k(A, k + i(A,k)) + \beta(1-\delta) \int V_k(A', (1-\delta)(k + i(A,k)))Q(\mathrm{d}A', A) \quad (26)$$

which can be decomposed as:

$$V_k(A,k) \quad = \quad \begin{cases} p & if k < s(A) \\ R_k(A,k) + \beta(1-\delta) \int V_k(A', (1-\delta)k)F(\mathrm{d}A', A) & if k > s(A) \end{cases} \quad (27)$$

Proof It is necessary to prove that the sequence of functions $W_k^j(A,k)$ converges uniformly to a continuous function $\tilde{W}_k(A,k)$ for every $(A,k,j) \in \mathcal{A} \times \mathcal{K} \times \mathbb{N}$. Solving W^j recursively, the following expression can be computed:

$$\begin{aligned} & W^{J+1}(A_0,k) - W^J(A_0,k) \\ = \quad & \beta^{J+1}(\textstyle\prod_{l=1}^J Q([\underline{A}, s^{l,-1}(k(1-\delta)^l)], A_{l-1})) \int_{\underline{A}}^{s^{J+1,-1}(k(1-\delta)^{J+1})} R(A_{J+1}, k(1-\delta)^{J+1}) \\ + \quad & \int_{s^{J+1,-1}(k(1-\delta)^{J+1})}^{\bar{A}} [W^{J+1}(A_{J+1}, S^{J+1}(A_{J+1})) - pS^{J+1}(A_{J+1}) + pk(1-\delta)^{J+1} - F]Q(\mathrm{d}A_{J+1}, A_J) \end{aligned}$$

Differenting with respect to k:

$$
\begin{aligned}
& W_k^{J+1}(A_0,k) - W_k^J(A_0,k) \\
= \;& (\beta(1-\delta))^{J+1}(\prod_{l=1}^{J} Q([\underline{A}, s^{l,-1}(k(1-\delta)^l)], A_{l-1})) \cdot [\\
& \int_{\underline{A}}^{s^{J+1,-1}(k(1-\delta)^{J+1})} R_k(A_{J+1}, k(1-\delta)^{J+1}) + \int_{s^{J+1,-1}(k(1-\delta)^{J+1})}^{\bar{A}} p \\
+ \;& \prod_{i=1}^{J} \frac{q(s^{i,-1}(k(1-\delta)^i), A_{i-1})}{Q([\underline{A}, s^{l,-1}(k(1-\delta)^l)], A_{l-1})} \cdot (\int_{\underline{A}}^{s^{J+1,-1}(k(1-\delta)^{J+1})} R(A_{J+1}, k(1-\delta)^{J+1}) \\
+ \;& \int_{s^{J+1,-1}(k(1-\delta)^{J+1})}^{\bar{A}} [W(A_{J+1}, S^{J+1}(A_{J+1})) - pS^{J+1}(A_{J+1}) + pk(1-\delta)^{J+1} - F])\} Q(\mathrm{d}A_{J+1}, A_J) \\
+ \;& (s_k^{J+1,-1}(k(1-\delta)^{J+1}) q(s^{J+1,-1}(k(1-\delta)^{J+1}), A_J)) \cdot \big[R(s^{J+1,-1}(k(1-\delta)^{J+1}), k(1-\delta)^{J+1}) - \\
& W(s^{J+1,-1}(k(1-\delta)^{J+1}), S^{J+1}(s^{J+1,-1}(k(1-\delta)^{J+1}))) - pS^{J+1}(s^{J+1,-1}(k(1-\delta)^{J+1})) + pk(1-\delta)^{J+1}
\end{aligned}
\tag{29}
$$

From the preceding expression and the boundedness of $W_k^j(A_0,k)$ for every $(A,k,j) \in \mathcal{A} \times \mathcal{K} \times \mathbb{N}$, the sequence $\left\{W_k^j(A_0,k)\right\}_{j=0}^{\infty}$ converges uniformly. From the point-wise convergence of $\left\{W^j(A_0,k)\right\}_{j=0}^{\infty}$, the result follows.

□

The necessary condition satisfied by the optimal stock of capital when there is an adjustment is:

$$R_k(A, S(A)) + \beta(1-\delta)\int V_k(A', (1-\delta)S(A))Q(\mathrm{d}A', A) = p$$

Using the strict concavity of R and the concavity of the integral, it follows that there exist a single value of k that satisfies the preceding equation. Since O a single-valued upper-hemi continuous correspondence, it is continuous. It is then straightforward to see that $S(A)$ and $s(A)$ are continuous.

Applying the implicit function theorem to the first-order condition and to the definition of $s(A)$ gives:

$$S'(A) = -\frac{W_{kA}(A, S(A))}{W_{kk}(A, S(A))} > 0 \tag{30}$$

$$s'(A) = -\frac{W_A(A, s(A)) - W_A(A, S(A))}{W_k(A, s(A)) - p} > 0 \tag{31}$$