

Optimal investment policy with fixed adjustment costs and complete irreversibility

Nicolas Roys

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP29/07

Optimal Investment Policy with Fixed Adjustment Costs and Complete Irreversibility

Nicolas Roys *

CeMMAP and Université Paris 1

November 18, 2007

Abstract

We develop and solve analytically an investment model with fixed adjustment costs and complete irreversibility that reproduces observed investment dynamics at the micro-level. We impose a minimal set of restrictions on technology and uncertainty. Most of the results duplicate or generalize earlier findings that have been established either by simulations or under counterfactual assumptions.

JEL: C61, D21 and E2

Keywords: Investment, Adjustment Costs, Irreversibility, Dynamic Programming

1 Introduction

It is well-known that investment is lumpy at the plant-level with long periods of inactivity punctuated by infrequent and large adjustments. For example, [17] document that more than half of continuing and large U.S. manufacturing establishments exhibit capital growth close to 50 percent in a single year and that 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 replicate observed microeconomic investment patterns satisfactorily, it is necessary to consider fixed adjustment costs and irreversibility.¹ The former are independent of the level of investment and incurred whenever investment is nonzero. A

*nicolas_r@ifs.org.uk. The author is grateful to Jean-Marc Robin, Cuong Le Van, François Gourio, Simon Gilchrist, Russell W. Cooper and Daniel Hamermesh for helpful discussions and Ben Skrainka for his corrections. The author thanks seminar and conference participants at the 62nd ESEM, the 11th Real Options Conference at Berkeley, Université Paris 1 and University of Texas at Austin. All errors are our own.

¹ [24] and [26] survey the literature on adjustment costs. While still widely used for their tractability, convex adjustment costs models imply a smooth and continual adjustment to shocks that is at odds with the data.

plant could avoid them by not investing at all. The latter depict a resale price of capital less than its current acquisition price. Capital specificity, *lemons effects*, and market thinness are important causes of irreversibility.

This paper analytically characterizes investment dynamics with fixed adjustment costs and complete irreversibility under a minimal set of restrictions on technology and uncertainty. In this way our analysis covers a wide class of problems.

An important restriction we impose throughout the paper is complete irreversibility: installed capital is valuable only to the extent that it is used in production. This approach describes many observed situations well since plants sell capital very rarely. Scarce episodes of negative investment are usually interpreted as depreciation and obsolescence of existing capital stock. It may also concern distressed firms (See [30] for a concrete example). In this case, capital sales and investments are two separate activities that do not interact. Only the latter is part of regular productive activities. In most empirical analysis, the dynamics of dis-investment is either estimated separately (See [3] for example) or neglected.

The following definition gives a formal derivation of observed investment policy and introduces notation that will be defined more precisely below.

Definition 1 *An investment policy, denoted i , is a state-dependent (S, s) policy if there exists a target function and a threshold function, denoted S and s , that do not depend on past levels 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 denotes the state of the world

Under such a policy, given the state of the world A , inaction is optimal when past level of capital (k) is above the threshold (s). Otherwise an investment episode occurs and may be important if the target (S) is sufficiently different from k .

Several investment models establish the optimality of state-dependent (S, s) policies either by simulations or under counterfactual assumptions. We believe that another investment model is needed for at least four reasons. First, a substantive difference between our approach and the theoretical literature is that we consider a true fixed-cost in the sense that it does not depend on the scale of operation. It allows the range of inaction to be wider for smaller plants. The main motivation is the stylised fact that smaller plants are characterized by substantially more intermittent investment. For example, [27] found that the frequency of zero investment for small plants is more than two times than for large plants.² Their empirical analysis supports the existence of a purely fixed cost component, unrelated to plant size.³ We

²This is also true for labor. [28] found that the frequency of episodes characterized by no employment changes decreases markedly with size. In principle, our model could be used in a labor context but the absence of job destruction would be inappropriate.

³Financial constraints are not the most likely explanation: they found a similar pattern in small-plants belonging to multi-plants large firms and also when they split the sample according to criteria correlated with the probability of financing constraints. Indivisibilities may also play a role.

are not aware of another characterization of optimal investment policy in this framework. [18] studies a special case where $S = s^4$ which arise when adjustment costs are linear but not fixed. In [2], there is a fixed cost and inaction may be optimal, but no lumpiness in the sense that $S = s$ because of their continuous-time setting, in which costs are proportional to the duration of the investment flow. In [14] and [13], a scale-dependent fixed cost reflects forgone profits due to production loss associated with installation of capital and is such that fixed investment costs are proportional to the existing capital stock. This property implies that fixed costs do not become irrelevant as the firm grows.

Second, our results hold for a wide-class of technologies, with very few restrictions on the nature of the uncertainty confronting the plant. In comparison, the existing literature always assumes normality. However in many economic situations, fat tails and asymmetries suggest that normality is too restrictive. Also, most papers use a geometric Brownian motion for tractability. Recent development of tests for units roots using panel data (See for example [22]) demonstrate that in most situations the presence of a unit root is rejected. While the optimal policy is stated without functional forms assumptions in [2], the rest of their analysis and all the associated papers use specific functions and homogeneity assumptions on technology. As stated by [4]: *estimation of production functions has a long history in applied economics ... Unfortunately, this history cannot be deemed an unqualified success, as many of the econometric problems that hampered early estimation are still an issue today.* Characterizing optimal policies without relying on any specific functional form is desirable.

Third, despite the non-differentiability and the discontinuity of the one-period profit function, the value function is continuous and differentiable almost everywhere except at the threshold $s(A)$ where it admits unequal left and right derivatives. Many have left the question of differentiability to further research and we are not aware of another proof in a dynamic programming model with fixed cost. An analytical expression for the marginal value of capital is derived and interpreted. The optimal investment policy depends simultaneously on the marginal value of capital and the total value of capital.

Fourth, analytical results have only been given in continuous-time.⁵ Yet, there is growing research that demonstrates the advantages of working in discrete-time (See [11] for an extensive discussion).

To summarize, this paper provides an explanation for observed investment policy under either less restrictive or more realistic assumptions that has been done before. Why is this important? Aside generalizing earlier findings, we highlight implications for three strands of literature. First, it is related to the empirical literature on investment dynamics.⁶ Since observations are available at equally spaced moments in time, the analysis is naturally in discrete-time. Reduced-form empirical papers customarily argue that an (S, s) policy is optimal. In structural models, sev-

⁴This paper deals with positive and negative demand for inputs and an arbitrary number of factors.

⁵ [14] is an exception: their model considers a very specific shock process, scale-dependent fixed costs, and imposes homogeneity assumptions.

⁶See [10] for a recent survey.

eral studies discuss the optimality of inaction and jumps in the investment process using numerical procedures. But both lack formal proof: either the (S, s) policy is observed in numerical examples or they refer the reader to theoretical paper done in continuous-time. Our main proposition, which establishes that the optimal policy is of the (S, s) form, can serve as a theoretical reference for empirical papers.

Also, the characterization of the optimal investment policy helps the identification strategy. In methods relying on user-designated moments, it helps understanding the link between those moments and structural parameters. For example, the serial-correlation in investment rates is a widely-used moment. Yet, we show that it does not permit to identify non-convex adjustment costs.

Some papers attempt to estimate variants of our framework using maximum likelihood and share the common feature of reducing the problem to a dynamic discrete choice model. For example, [31] estimates a labor demand model with fixed costs. The continuous decision (how much to invest) is estimated non-parametrically and it is the discrete decision (whether to adjust or not) that identify the structural parameters. By full-characterizing the optimal policy, we can form a likelihood for the investment choice using a selection model: one equation for whether or not to invest and one equation for how much to invest.⁷ Recovering the structural parameters would then involve both continuous and discrete control variables which may avoid important information losses.

Second, knowledge of the structure of the value functions aid the design of approximation algorithms to estimate them. Using value function iteration, we show that restricting the search to policies of the (S, s) form reduces drastically the computational-time required to solve the model.

Lastly, our results have implications for the literature on investment-cash flows sensitivities which uses [25] as a benchmark. The latter assumes a perfectly competitive firm, a constant-returns-to-scale production function and convex investment adjustment costs that are homogeneous of degree one in investment and capital. [25]'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 the replacement cost of 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 . These findings are usually interpreted as the presence of financial frictions. Recent papers have challenged this interpretation.⁸ They shows that 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 literature, an solution (actually infeasible) would be to replace average q (observable) by marginal q (unobservable) in regressions. We show that, without financing constraints, the optimal investment policy depends simultaneously on marginal and average q . It follows that focusing on these reduced-form investment equation is quite problematic.

The rest of the paper is organized as follows. Section 2 presents the model.

⁷ [6] approach incorporates continuous and discrete choices and could be used to estimate the value function which belong to the set of covariates in both equation.

⁸A few examples are [21], [16] and [7].

Section 3 derives the optimal decision rule. Section 4 establishes the differentiability of the value function and re-formulates investment policy in terms Tobin's average and marginal q. Section 5 concludes. All proofs are in the Appendix.

2 The Model

In this section, we introduce the model and impose assumptions under which the (S, s) policy is optimal. For the sake of notational simplicity, we consider a version of the model where parameters are constant over time, the horizon is infinite and the stochastic process is stationary and univariate. Section 2.3 relaxes some assumptions.

2.1 Assumptions

Time is discrete and indexed by t . At each period, the plant decides to invest or not ($i_t \geq 0$) over an infinite time horizon. Her decision depends on k_t the level of capital inherited from the previous period, and A_t an observable and univariate shock to the profitability of the plant at the beginning of period t . The latter combines product demand, productivity of inputs, and prices of flexible factors (wages, energy prices, etc...) conditions.⁹ The one-period profit function is:

$$\pi(A, k, i) = R(A, k + i) - C(i) \quad (1)$$

$R(A, k)$ represents reduced-form profits and incorporates the optimal choice of flexible factors. It is based on underlying production and cost functions where all flexible factors of production (labor, working time, etc.) have been optimized out (See Section 3.4 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, one assumes that investment becomes productive with one period lag. Since it would complicate the notation without reversing the results,¹⁰ assume it is immediately productive.

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

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

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 $\lim_{i \rightarrow 0^+} C(i) > 0$. We present the timing of the model and then the technical assumptions that will be used in the proofs.

Assumption 1 *Timing of the model and State Space*

⁹See [19] for an empirical attempt to decompose these shocks.

¹⁰The 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 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.

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 decides to invest or not ($i_t \geq 0$)
3. The level of capital at the start of the next period, $t + 1$, is:

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

where δ is a positive and constant depreciation rate.¹¹

4. New value of profitability are drawn from a Markov transition function $Z : \mathbf{A} \times \mathcal{A} \rightarrow [0, 1]$ where $(\mathcal{A}, \mathbf{A})$ is a measurable space.

Assumption 2 The plant is risk-neutral and discounts future profits at a constant rate $\beta \in (0, 1)$.

Assumption 3 $Z(A_{t+1}, \cdot)$ is continuous and stochastically increasing in the first-order stochastic dominance sense for every $A_{t+1} \in \mathcal{A}$.

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

Under the assumption of risk-neutrality, the plant's objective is to maximize its expected net present value which is the discounted sum of one-period profits net of adjustment costs. A follows an exogenous and stationary Markov process. Technology improvements and adoption can easily be introduced in the model (See [12] among others). If the latter render the problem nonstationary, [9] show how nonstationary models can be reduced to stationary ones by appropriate reformulation. While exogeneity may appear as an important restriction on the learning process, it is consistent with the empirical literature, at least for some industries.¹² An empirically relevant interpretation of Assumption 4 is that at the beginning of her life, the manager draws a permanent level of profitability from which actual profitability can deviate according to a Markovian process. Z can be thought-of as conditional on the initial conditions. Furthermore, in Section 2.3, we shows that the optimality of an (S, s) policy holds with almost no restrictions on the form of uncertainty.

From Assumption 3, 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. The transition function $Z(A', A)$ satisfies the Feller property: it guarantees that the expectation function, used later in 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.

¹¹To be consistent with the assumption that investment is immediately productive, depreciation is proportional to both accumulated capital and investment.

¹² [29] find that manufacturer sales appears to be Markovian whereas retailer sales do not. Using a sample of Texas Bars, [1] find that Non-Markovian dynamics arise from permanent and unobservable differences across entrepreneurs' choices of their firms' intended scales

The profit function includes flows payments on capital such as maintenance costs or commitments associated with the purchase of capital. Combined with decreasing returns and/or some degree of monopoly power, $R(A, \cdot)$ is concave in k . The latter may appear to strong in that it is imposed for every level of profitability. [5] presents a generalization of stochastic technology where although the production process exhibits decreasing returns-to-scale in some average sense, it may enjoy increasing returns for some favorable resolutions of the uncertainty. At the cost of further notational and regularity assumptions, the results of the present paper can be extended to such a situation.

Finally, to save on notation, we ruled out the existence of convex adjustment costs. This assumption is very convenient but unnecessary for our results to hold. And as discussed below, decreasing-returns and convex adjustment costs impact similarly investment patterns.

2.2 The Dynamic Programming Problem

Given the law of motion of capital $k_{t+1} = (1-\delta)(k_t + i_t)$ and k_0 , a manager chooses the sequence of investment $\{i_t\}_{t=0}^{\infty}$ to maximize the present discounted value of current and future profits:

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

In Appendix A, we show that the objective function defined in (4) is well-defined and the capital state space is compact. The latter is not ad-hoc assumption: the capital stock can not take negative values and investing when capital is above a particular value (denoted \tilde{k}) can not be optimal. We now write the model as a dynamic programming problem and its standard recursive formulation.

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

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

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

$$V(A, k) = \sup_{0 \leq i \leq \tilde{k} - k} [W(A, k + i) - C(i)] \quad (6)$$

where

$$W(A, k) \equiv R(A, k) + \beta \int_{\mathcal{A}} V(A', k(1-\delta)) Z(dA', A) \quad (7)$$

W can be interpreted as the value of the plant when investing today is not allowed and the optimal policy is followed in all future periods. While the 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 uniqueness of a solution Bellman's equation.

More importantly the value function is continuous despite the discontinuity in the one-period profit function. Note that Proposition 1 requires few assumptions about the stochastic process.

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

Proof see Appendix A □

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

$$O(A, k) = \left\{ k' \in [k, \tilde{k}] : V(A, k) = W(A, k') - C(k' - k) \right\} \quad (8)$$

Next Corollary is a consequence of the Maximum Theorem,

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

2.3 A More General Framework

We impose assumptions that are strong enough to be tractable for empirical work. However, it should be clear that almost no restrictions have to be imposed on the form or nature of the uncertainty confronting the plant. The model allows non-stationary and multivariate shocks, a horizon of arbitrary length and time-varying parameters. Consider the probability space $(\mathcal{A}, \mathcal{F}, P)$ and filtration $\mathbb{F} = \bigcup_t \mathcal{F}_t$ where \mathcal{F}_t represents the information available at time t . Let $A \in \mathcal{A}$ be the state of the world. The one-period profit function in period t is denoted $R_t(A, k)$. For each $t \in \{1, \dots, T\}$, R_t is jointly measurable, $R_t(\cdot, k)$ is \mathcal{F}_t -measurable for each $k \in \mathbb{R}^+$ and $R_t(A, \cdot)$ is concave for each $A \in \mathcal{A}$. 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 necessary to assume that the discounted fixed cost sequences $\{\beta^{t-1} F_t\}_{t=1}^{T+1}$ are non-increasing in t , which holds trivially when these costs are fixed over-time.

Under standard regularity conditions, for all $t \in \{1, \dots, T\}$, the value function V_t is jointly measurable and is given by the solution to Bellman's equation:

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

The final-period value function V_{T+1} is assumed jointly measurable and $V_{T+1}(A, \cdot)$ is F_{T+1} -concave (cf. infra) for each $A \in \mathcal{A}$. It is straightforward to see that most of the result holds in the more general framework presented here.

3 Optimal Decision Rule

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 F -concavity introduced by [32] can be used to characterize the optimal decision rule.

3.1 Mathematical Background

There is two equivalent ways to define F -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 $\lambda \in [0, 1]$,

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

The Next Proposition's proof can be found in [32], except the last property whose proof is available from the author upon request.

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 $g = \lim_{n \rightarrow \infty} g_n$ 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, where $F \geq 0$, then the function

$$g(x) = \max_{y \geq x} \{f(y) - FI\{y > x\}\}$$

is also F -concave

3.2 Optimality of an (S, s) Policy

Essentially, we tailor the arguments for a class of commodity price speculation problems in [23] to derive the optimality of an (S, s) investment policy.

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:

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

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

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

$$s(A) = \min \left(k \in [0, \tilde{k}] : W(A, k) \geq W(A, S(A)) - p[S(A) - k] - F \right) \quad (11)$$

The value function V can be expressed as:

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

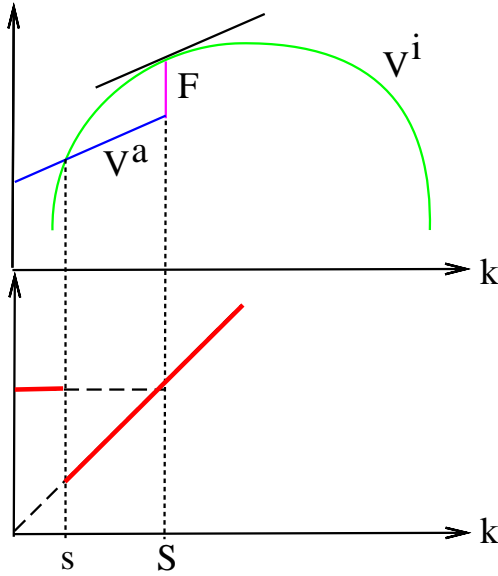


Figure 1: Value Function and Optimal Decision Rule

Proof See Appendix C □

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 $s(A)$, the optimization problem is separated in two parts: choosing an optimal level of investment ignoring the fixed cost,¹³ 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.

To handle the case where there are multiple maximizing values, the target $S(A)$ is defined as the smallest value of the policy correspondence $O(A)$.

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.

For illustrative purposes, we 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 levels, the gap between the value of inaction and the value of an (optimal) adjustment is exactly equals to the fixed adjustment cost. Without linear adjustment costs, the value of ordering would equate the value of inaction evaluated at the target level minus the fixed cost.

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

¹³It does not appear in the definition of $S(A)$. Yet it impacts on $S(A)$ through its effect on the value function.

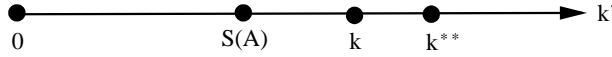


Figure 2: Value Function and Optimal Decision Rule

the slope of the blue curve ¹⁴. 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 less than their respective costs.

At the level of previous capital in which the plant is indifferent between adjustment and non-adjustment, the level of capital chosen, if the plant adjusts, is larger than previous capital. Irreversibility alone does not produce this phenomena. This is due to the presence of a fixed cost the plant aims to avoid incurring too frequently. Furthermore 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 and the latter effect vanishes.

Next corollary directly follows from Proposition 3. We discuss informally its proof because it gives insights onto investment behavior.

Corollary 2 *k is not an argument of the target function S .*

Corollary 2 says that as long as investment is strictly positive, 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, 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: $k' \in]k, \tilde{k}]$. Indeed, this is a problem only if $S(A) < k$. A situation like the one in the Figure 3.2 is conceivable: if a negative investment level were allowed, it would be the arg max of the objective function). Because it is not, the plant prefers to invest (and choose the optimal level of capital belonging to the admissible state space, 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:

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

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

This property is interesting for empirical analysis. To evaluate the empirical plausibility of investment models with non-convexities in adjustment costs, the serial-correlation of investment rates is a widely-used moment in the empirical literature.

¹⁴The differentiability of the value function is established in Section 4.1

Standard Algorithm	547.3 sec
(S,s) Algorithm	6.7 sec

Table 1: Computational-Time

But given the preceding result, it is not an appropriate moment to identify irreversibility and fixed costs. Then, it may reflect either persistence of the shock process, convex adjustment costs or a plant fixed-effect.

3.3 Numerical Solution

Even in times of increasing computer-power, a potential value of our approach is that it allows to reduce the computational costs. Under the assumptions made, when one solves the model numerically, he can restricts his attention to policies of the (S, s) form. To evaluate those gains, we solve the model using value function iteration. Since the algorithm is well-known, we do not discuss the details.

We first use the standard version of the algorithm: at each iteration of the Bellman equation, for each point of the productivity grid, the value of choosing a particular point in the capital grid is evaluated and the maximum over all those points is chosen as the optimal policy. We then use the structure of the value function to solve the model. Using Corollary 2, at each iteration of the Bellman equation, for each point of the productivity grid, it is sufficient to look at any point of the capital state space to find the optimal target (S) . Hence during the search of the optimal value of capital, the number of operations is divided by the size of the capital grid. To compute the value function, we use Proposition 3.

Cash-flows are specified as $R(A, k) = Ak^\alpha - rk$. r is such that the steady-state capital be 700. $\log(A)$ follows an $AR(1)$ with persistence ρ and standard deviation σ . We use a standard calibration: $\alpha = 0.6$, $\beta = 0.98$, $p = 0.02$, $F = 3.5$, $\rho = 0.9$ and $\sigma = 0.3$. Table 3.3 reports the computational-time for a reasonable sized grid (300 points for capital and 20 points for productivity) and convergence criteria ($1e - 6$).

Taking into account the structure of the optimal policy lead to a drastic reduction of the computational-time that is divided by about 80. The number of iterations before convergence is the same. The computational gain comes for the way the value function is solved. Our findings are limited to value function iteration which is a natural benchmark. We are optimistic that restricting the search to policies of the (S, s) could be shown to alleviate the computational burden for alternative solution methods, but we leave this for future research.

3.4 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 plant chooses its optimal level at no cost, past level of these inputs will not impact their current level. However, it is necessary to assume that after having made the optimal choices for 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 realistic¹⁵, 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 plant 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 profit-function and assume 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 the profit function has positive (negative) derivatives 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 [15]. 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 shocks. But if the plant hires new workers then working-time eventually stays at its regular-level.

Building on the preceding example, the general formula for reduced-form profits is:

$$R(A, k) = \int \max_h \tilde{R}(A, k, h, \mu) F_\mu(d\mu) \quad (13)$$

where h is an arbitrary vector of flexible inputs, μ is an *i.i.d* shock with cdf F_μ .

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, we show that the value function is differentiable *a.e.* except at the threshold level where it admits unequals left and right derivatives. It holds despite the non-differentiability and the discontinuity of the one-period profit function.

4.1 Differentiability of the Value Function

To make possible the analysis in terms of marginal and average q , we henceforth assume that the reduced-form profit function is differentiable. Obviously differentiability of the one-period profit function is necessary to expect differentiability of the value function. Supermodularity says that the marginal productivity of capital is increasing in profitability.

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

¹⁵Notably, to avoid a deterministic relationship between the two factors.

Proposition 4 $V(A, \cdot)$ is differentiable in k almost everywhere on the domain $\mathcal{K} \setminus s(A)$ for every $A \in \mathcal{A}$. The partial derivative obeys the equation:

$$V_k(A, k) = R_k(A, k + i(A, k)) + \beta(1 - \delta) \int V_k(A', (1 - \delta)(k + i(A, k))) Z(dA', A) \quad (14)$$

which can be decomposed as:

$$V_k(A, k) = \begin{cases} p & \text{if } k < s(A) \\ R_k(A, k) + \beta(1 - \delta) \int V_k(A', (1 - \delta)k) Z(dA', A) & \text{if } k > s(A) \end{cases} \quad (15)$$

Corollary 3 For every $A \in \mathcal{A}$, the target $S(A)$ satisfies

$$R_k(A, S(A)) + \beta(1 - \delta) \int V_k(A', S(A)(1 - \delta)) Z(dA', A) = p \quad (16)$$

Corollary 4 The functions S and s are non-decreasing

Proof See Appendix D □

Despite the non-differentiability and the discontinuity of the one-period profit function, Proposition 4 establishes the differentiability of the value function almost everywhere except at the threshold where it admits unequal left and right derivatives. The thrust of our proof is to consider a finite-version of the model and to show that the sequence of derivatives associated with each finite horizon converge uniformly. From the convergence of the value function, the result follows.

The interpretation of Corollary 3 is the usual one: if the plant decides to invest, it equates marginal gains from adjustment to its marginal costs. Note that the fixed cost does not appear explicitly in first-order condition. However, it impacts the target indirectly through its effects on the value function.

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

To compute the derivative with respect to capital of the value function, we solve recursively the function W . Define $a(k) = \min\{A \in \mathcal{A} | W(A, k) \geq W(A, S(A)) - p[S(A) - k] - F\}$. Since $\lim_{j \rightarrow \infty} \beta^j \|W^j\| = 0$, it follows that:

$$\begin{aligned} & W(A_0, k) \\ = & R(A_0, k) + \beta \int_{\underline{A}}^{a(k(1-\delta))} R(A_1, k(1-\delta)) Z(dA_1, A_0) \\ & + \beta \int_{a(k(1-\delta))}^{\bar{A}} [W(A_1, S(A_1)) - pS(A_1) + pk(1-\delta) - F] Z(dA_1, A_0) \\ & + \sum_{j=2}^{\infty} \beta^j \left(\prod_{l=1}^{j-1} Z([\underline{A}, a(k(1-\delta)^l)], A_{l-1}) \right) \left(\int_{\underline{A}}^{a(k(1-\delta)^j)} R(A_j, k(1-\delta)^j) Z(dA_j, A_{j-1}) \right. \\ & \left. + \int_{a(k(1-\delta)^j)}^{\bar{A}} [W(A_j, S(A_j)) - pS(A_j) + pk(1-\delta)^j - F] Z(dA_j, A_{j-1}) \right) \quad (17) \end{aligned}$$

An analytic expression for the marginal value of capital is given in Appendix D. It can be decomposed 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 plant adjusts and chooses a new level of capital, the target, which has been proven independent of past values of capital. Consequently 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 take this into account. An increase in capital decreases the probability of an adjustment in the future. There is a loss of the future marginal gains to adjust. Yet by definition, if the plant actually adjusts both total gains of adjustment and the marginal gains of adjustment are larger than their respective costs. The third effect is positive. By increasing capital, the plant reduces its probabilities of adjusting in the future which increase the number of periods where the accumulated capital will have an impact on the marginal value of the plant.

4.2 Investment, Marginal and Average Value of Capital

The optimal policy can be formulated in terms of the marginal and the average value of capital. Define the average value of capital, also known as Tobin's q , as $Q(A, k) = \frac{W(A, k)}{pk}$. Define the marginal value of capital, also known as marginal q , as $q(A, k) = \frac{W_k(A, k)}{p}$. It is implicit in this subsection that $p > 0$. The case $p = 0$ is left to the reader. Proofs in this subsection follow from preceding results and are therefore omitted:

Corollary 5 *For every $A \in \mathcal{A}$, $S(A)$ is such that $q(A, S(A)) = 1$*

So as long as investment is positive, investment only depends on marginal q . This holds whether fixed costs are positive or not. Yet, fixed costs impact on marginal q dynamics:

Proposition 5 *If $F = 0$, for every $(A, k) \in \mathcal{A} \times \mathcal{K}$, $q(A, k + i(A, k)) \leq 1$ and*

$$q(A, k + i(A, k)) = \begin{cases} q(A, k) & \text{if } q(A, k) \leq 1 \\ 1 & \text{if } q(A, k) > 1 \end{cases} \quad (18)$$

Proposition 6 *If $F > 0$, for every $(A, k) \in \mathcal{A} \times \mathcal{K}$*

$$q(A, k + i(A, k)) = \begin{cases} q(A, k) & \text{if } Q(A, k) > 1 + (Q(A, S(A)) - 1) \frac{S(A)}{k} - \frac{F}{pk} \\ 1 & \text{otherwise} \end{cases} \quad (19)$$

Corollary 6 *If $F > 0$, for every $A \in \mathcal{A}$, there exists $k \in (s(A), S(A))$ such that $q(A, k + i(A, k)) > 1$*

Proposition 5 is a standard result from the investment literature: it is profitable to invest if and only if marginal q is larger than 1. Since marginal costs of adjusting are constant, an inaction region appears even if there are no fixed costs.

The analog of Proposition 5 exists in investment models with constant return-to-scale and convex adjustment costs. There the latter prevent investment from going to infinity. Decreasing return-to-scale and/or market power (Assumption 4) has the same effect here. A crucial difference is that convex adjustment costs generate a smooth and continuous adjustment to shocks for which there is little evidence in the data while decreasing return-to-scale and/or market power do not.

With a positive fixed cost, the probability of an adjustment depends only on Tobin's average q . It entirely explains the decision to adjust or not. Then, the optimal amount of investment is entirely determined by Tobin's marginal q , which is a sufficient statistic for the marginal gains of an adjustment. With a positive fixed cost, in some state, the plant may not wish to adjust its capital stock even if the marginal q is larger than 1. This is in contrast with the result of the literature which state that marginal q always remains below 1 since when it cross this threshold, marginal gains of an adjustment are larger than marginal costs and the plant adjusts its capital stock to its desired level. Yet because of the fixed cost, total gains from adjustment have to be larger than total costs for the plant to adjust. Such a situation appears when the beginning of the period capital lies above the threshold $s(A)$ but below the target $S(A)$.

Because of decreasing returns to investment and/or market power, marginal and average q are not the same (See [25]). While we present an analytic expression for both, respectively, in Equation 17 and Appendix D, we are not able to characterize the way the two expressions differ. Nevertheless, cash-flows are correlated with both. So the significance of cash-flows in standard reduced-form regressions may not reflect finance constraints. The link between cash-flows and investment is thoroughly investigated by [16]: *the response of investment to cash-flow indicates the additional information, relative to average q , contained in this variable for forecasting future profitability.* We confirm their result analytically while their analysis uses simulation.

In Equation 19, the fixed cost is divided by the level of the capital stock which captures the level of operation. It follows that fixed costs matter more for small plants. Thus, the model delivers a inaction region wider for small plants in agreement with recent empirical evidence (cf. *supra*).

5 Conclusion

This paper proves the high degree of generality of (S, s) policies in an model of investment with a fixed cost independent of plant-size. We establish uniqueness, continuity and differentiability (except at the threshold) of the value function. The optimal policy is characterized analytically and depends simultaneously on marginal and average q .

There are several immediate extensions to this paper. We do not analyse disinvestment episodes. This should not be an assumption but rather a feature of the model. Yet, it creates technical difficulties. The value function is not F -concave and

the optimal policy is no longer guaranteed to be of the (S, s) form. Also, capital is not an homogeneous good and modelling capital heterogeneity is a relatively unexplored line of research. [20] considers a form of K -convexity in \mathfrak{R}^n . Yet, this notion persists after a dynamic programming iteration only in very special cases. We leave this for future research.

References

- [1] Jaap H. Abbring and Jeffrey R. Campbell. A firm's first year. Tinbergen Institute Discussion Papers 05-046/3, Tinbergen Institute, May 2005.
- [2] Andrew B Abel and Janice C Eberly. A unified model of investment under uncertainty. *American Economic Review*, 84(5):1369–84, December 1994.
- [3] Andrew B. Abel and Janice C. Eberly. Investment, q with fixed costs: An empirical analysis. mimeo, 2002.
- [4] Daniel Akerberg, Lanier Benkard, Steve Berry, and Ariel Pakes. *Econometric Tools for Analyzing Market Outcomes*, volume 6, chapter Handbook of Econometrics. 2007.
- [5] Rabah Amir. A new look at optimal growth under uncertainty. *Journal of Economic Dynamics and Control*, 22(1):67–86, November 1997.
- [6] Patrick Bajari, C. Lanier Benkard, and Jonathan Levin. Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370, 09 2007.
- [7] Nihal Bayraktar, Plutarchos Sakellaris, and Philip Vermeulen. Real versus financial frictions to capital investment. Working Paper Series 566, European Central Bank, December 2005.
- [8] Lawrence M. Benveniste and Jose A. Scheinkman. On the differentiability of the value function in dynamic models of economics. *Econometrica*, 47(3):727–32, May 1979.
- [9] Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic Optimal Control: The Discrete Time Case*. 1978.
- [10] Stephen Bond and John Van Reenen. *Microeconomic Models of Investment and Employment*, volume 6 of *Handbook of Econometrics*. 2007.
- [11] Svetlana Boyarchenko and Sergei Levendorskii. Practical guide to real options in discrete time. *International Economic Review*, 48(1), 2007.
- [12] Thomas Buettner. R_d and the dynamics of productivity. mimeo, London School of Economics, 2004.
- [13] Ricardo J. Caballero and Eduardo M. R. A. Engel. Explaining investment dynamics in u.s. manufacturing: A generalized (s,s) approach. *Econometrica*, 67(4):783–826, July 1999.
- [14] Ricardo J. Caballero and John V. Leahy. Fixed costs: The demise of marginal q . NBER Working Paper 5508, National Bureau of Economic Research, Inc, March 1996.

- [15] Russel W. Cooper, John C. Haltiwanger, and Jonathan Willis. Dynamics of labor demand: Evidence from plant-level observations and aggregate implications. NBER Working Papers 10297, National Bureau of Economic Research, Inc, February 2004.
- [16] Russell Cooper and Joao Ejarque. Financial frictions and investment: Requiem in q . *Review of Economic Dynamics*, 6(4):710–728, October 2003.
- [17] Mark E. Doms and Timothy Dunne. Capital adjustment patterns in manufacturing plants. *Review of Economic Dynamics*, 1(2):409–429, April 1998.
- [18] Janice C. Eberly and Jan A. Van Mieghem. Multi-factor dynamic investment under uncertainty. *Journal of Economic Theory*, 75(2):345–387, August 1997.
- [19] Lucia Foster, John Haltiwanger, and Chad Syverson. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, forthcoming, 2007.
- [20] Guillermo Gallego and Suresh P. Sethi. k -convexity in \mathbb{R}^n . *Journal of Optimization Theory and Applications*, 127(1):71–88, October 2005.
- [21] Joao F. Gomes. Financing investment. *American Economic Review*, 91(5):1263–1285, December 2001.
- [22] Bronwyn H. Hall and Jacques Mairesse. Testing for unit roots in panel data: An exploration using real and simulated data. mimeo, UC Berkeley, 2003.
- [23] George Hall and John Rust. The (s, s) policy is an optimal trading strategy in a class of commodity price speculation problems. *Economic Theory*, 30(3):515–538, March 2007.
- [24] Daniel S Hamermesh and Gerard A Pfann. Adjustment costs in factor demand. *Journal of Economic Literature*, 34(3):1264–1292, September 1996.
- [25] Fumio Hayashi. Tobin’s marginal q and average q : A neoclassical interpretation. *Econometrica*, 50(1):213–24, January 1982.
- [26] Aubhik Khan and Julia K. Thomas. *Adjustment Costs*. The New Palgrave Dictionary of Economics. Palgrave Macmillan, 2006.
- [27] Oivind Nilsen and Fabio Schiantarelli. Zeros and lumps in investment: Empirical evidence on irreversibilities and nonconvexities. *The Review of Economics and Statistics*, 85(4):1021–1037, December 2003.
- [28] Oivind A. Nilsen, Kjell G. Salvanes, and Fabio Schiantarelli. Employment changes, the structure of adjustment costs, and plant size. *European Economic Review*, 51(3):577–598, April 2007.
- [29] Ariel Pakes and Richard Ericson. Empirical implications of alternative models of firm dynamics. *Journal of Economic Theory*, 79(1):1–45, March 1998.

- [30] Valerie A. Ramey and Matthew D. Shapiro. Displaced capital: A study of aerospace plant closings. *Journal of Political Economy*, 109(5):958–992, October 2001.
- [31] Paola Rota. Estimating labor demand with fixed costs. *International Economic Review*, 45(1):25–48, 02 2004.
- [32] Herbert E. Scarf. *The optimality of (S,s) policies in the dynamic inventory problem*. Mathematical Methods in the Social Sciences. Stanford University Press, 1960.
- [33] Manfred Schal. Conditions for optimality in dynamic programming and for the limit of n -stage optimal policies to be optimal. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 32(3):179–196, November 1975.
- [34] Nancy L Stokey, Robert Jr E Lucas, and Edward C Prescott. *Recursive Methods in Economic Dynamics*. Harvard University Press, 1989.
- [35] Donald M. Topkis. Minimizing a submodular function on a lattice. *Operations Research*, 26(2):305–321, 1978.

Appendices

All results should be understood to hold under the assumptions in the text.

A Proof of Proposition 1

Lemma 1 (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 solving a sequence of static decision problems. $R(A, \cdot)$ has a global maximizer for every $A \in \mathcal{A}$ and the expected discounted profits of such a plant is finite. Hence if the supremum exists, it must be finite. Normalizing the initial capital stock to 0, the sequence $i_t = 0$ for all $t \geq 0$ yields a return of 0. We found an investment sequence which yields a value greater than $-\infty$ which shows that the supremum exists. \square

Given Lemma 1, we can define the supremum function associated with (4), say V^* .

Lemma 2 There exists $\tilde{k} < \infty$ such that for every $t \in \mathbb{N}$, $k_t \leq \tilde{k}$.

Proof For every $A \in \mathcal{A}$, define $\tilde{k}(A)$ as $R(A, \tilde{k}) - F - p\tilde{k} + E[\sum_{t=1}^{\infty} \beta^t R(A_t, k_t + i_t)] = 0$. Let $\{k_t(A_t)\}_{t=0}^{\infty}$ be a sequence such that $\Pr[k_t(A_t) > \tilde{k}(A_t)] > 0$ for some $t > 0$. Then we 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 $\tilde{k}(A_1)$, $R(A_1, k_1) - F - pk_1 + E[\sum_{t=1}^{\infty} \beta^t R(A_t, k_t + i_t)] < 0$. Because $\sup_{\{i_t\}_{t=1}^{\infty}} E[\sum_{t=1}^{\infty} \beta^t \Pi(A_t, k_t, i_t)] \leq E[\sum_{t=1}^{\infty} \beta^t R(A_t, k_t + i_t)]$ and $V^* \geq 0$ (See Proof of Lemma 1), it follows:

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

Finally, let $\tilde{k} = \max_{A \in \mathcal{A}} \tilde{k}(A)$. \square

The following sets are defined under the usual sup-norm $\|\cdot\|$. Let $B_{\mathcal{AK}}$ be the set of bounded functions $V : \mathcal{A} \times \mathcal{K} \rightarrow \mathbf{R}$. Let $C_{\mathcal{AK}}$ be the set of continuous functions $V : \mathcal{A} \times \mathcal{K} \rightarrow \mathbf{R}$. The operator associated to Bellman equation 6: $T(V) = V$. The value function can be re-written as: $V = \max\{V^a, V^i\}$ where $V^a(A, k) = \sup_{0 \leq i \leq \tilde{k}-k} \{W(A, k+i) - C(i)\}$ and $V^i(A, k) = W(A, k)$. we show that the operator T maps $C_{\mathcal{AK}}$ into itself and has 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 Lemma 9.5 in [34] and Assumption 3, $V \in C_{\mathcal{AK}} \Rightarrow [\int_{\mathcal{A}} V(A', (1-\delta)(k+i)) Z(dA', A)] \in C_{\mathcal{AK}} \Rightarrow V^i \in C_{\mathcal{AK}}$. By the theorem of the maximum, $V^a(A, k) \in C_{\mathcal{AK}}$. Because the operator max is 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 3

Lemmas 3-6 and their proofs mimic analogous Lemmas in [23]. Therefore, proofs of Lemma 4 and 6 are omitted. Yet, we show that Lemma 4 is a straightforward consequence of a property of F -concave functions while [23] prove it in their framework only.

Define the operators Γ and Λ by:

$$\Gamma(W)(A, k) = \max_{0 \leq i \leq k-k} [W(A, k+i) - C(i)] \quad (20)$$

$$\Lambda(V)(A, k) = R(A, k) + \beta \int_{\mathcal{A}} V(A', k) Z(dA', A) \quad (21)$$

Let $\mathcal{F}_{\mathcal{A}\mathcal{K}}$ denote the class of functions $V(A, k)$ which are continuous¹⁶ in A and k , and F -concave as a function of $k \in \mathcal{K}$ for all $A \in \mathcal{A}$. Lemma 3 proves that if $W(A, k)$ is F -concave, so is V . Lemma 4 proves the optimality of an (S, s) policy. Lemmas 5 and 6 show that the composition operators $\Lambda \circ \Gamma$ and $\Gamma \circ \Lambda$ map $\mathcal{F}_{\mathcal{A}\mathcal{K}}$ into $\mathcal{F}_{\mathcal{A}\mathcal{K}}$.

Lemma 3 *If $W \in \mathcal{F}_{\mathcal{A}\mathcal{K}}$, then $V = \Gamma(W) \in \mathcal{F}_{\mathcal{A}\mathcal{K}}$.*

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

$$V(A, k) = \left(\max_{k' \geq k} [W(A, k') - pk' - FI\{k' > k\}] \right) + pk \quad (22)$$

where the first term is F -concave as a function of $k \in \mathcal{K}$ for all $A \in \mathcal{A}$ by Proposition 2.3 and so is the second term. \square

Lemma 4 *If $W(A, k) \in \mathcal{F}_{\mathcal{A}\mathcal{K}}$, then the solution to the functional equation (6) takes the form of a state-dependent (S, s) policy as described by Equations 9, 10 and 11.*

Lemma 5 $\Lambda \circ \Gamma$ maps $\mathcal{F}_{\mathcal{A}\mathcal{K}}$ into $\mathcal{F}_{\mathcal{A}\mathcal{K}}$.

Proof By Lemma 3 if $U \in \mathcal{F}_{\mathcal{K}\mathcal{C}}$, then $V = \Gamma(U) \in \mathcal{F}_{\mathcal{K}\mathcal{C}}$. By Lemma 4, there exist functions S and $s : \mathcal{A} \rightarrow \mathcal{K}$ satisfying $0 \leq s(A) \leq S(A) \leq \tilde{k}$ for which $V = \Gamma(U)$ can be represented as

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

Since positive linear combinations and point-wise limits of F -concave functions are F -concave, it follows that $\int_{\mathcal{A}} V(A', k(1 - \delta)) Z(dA', A) \in \mathcal{F}_{\mathcal{A}\mathcal{K}}$. Finally, $\Lambda \circ \Gamma(U)(A, k) = R(A, k) + \beta \int_{\mathcal{A}} V(A', k(1 - \delta)) Z(dA', A) \in \mathcal{F}_{\mathcal{A}\mathcal{K}}$. \square

Lemma 6 *The function $V(A, \cdot)$ and $W(A, \cdot)$ are F -concave function of $k \in \mathcal{K}$ for all $A \in \mathcal{A}$.*

¹⁶Continuity of V and W has already been proved.

C Proof of Proposition 4

Lemma 7 $V(A, k)$ is supermodular in (A, k)

Proof Let $S_{\mathcal{AK}}$ be the set of continuous, supermodular function $V : \mathcal{A} \times \mathcal{K} \rightarrow \mathbf{R}$ and such that $V(\cdot, k)$ is non-decreasing for each $k \in \mathcal{K}$. We show that T maps $S_{\mathcal{AK}}$ into itself. Consider $V \in S_{\mathcal{AK}}$. V can be written as $V(A, k') = \max_{k' \in [k, \bar{k}]} [W(A, k') - C(k' - k)]$. For a function on \mathbb{R}^2 , increasing differences is equivalent to supermodularity. We need to show that $\int V(A', k') Z(dA', A)$ 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 $k'_2 > k'_1$. Combined with $Z(A', \cdot)$ stochastically increasing, it follows:

$$\int [V(A', k'_2) - V(A', k'_1)] Z(dA', A_2) \geq \int [V(A', k'_2) - V(A', k'_1)] Z(dA', A_1)$$

which means that $\int V(A', k') Z(dA', A)$ 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') (as it is linear), $W(A, k') - C(k' - k)$ is supermodular in (A, k', k) . Maximizing over k' preserves supermodularity in the remaining variables (A, k) ([35]). Consequently, $V = T(V)$ is supermodular in (A, k) . \square

Lemma 8 S is a non-decreasing function.

Proof Following an approach simliar to Lemma 7 it can be shown that $W(A, k) - pk$ is supermodular in (A, k) . From [35], the set of optimal solutions is non-decreasing in A . $S(A)$ being the lowest element of the set, the result follows. \square

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

$$\begin{aligned} V^{j+1}(A, k) &= \max_{k^j \geq k} \left[R(A, k^j) - C(k^j - k) + \beta \int V^j(A', k^j(1 - \delta)) Z(dA', A) \right] \\ &= T^{j+1} V^0(A, k) \end{aligned}$$

where k^j is the optimal choice of capital for the $j + 1$ period problem. Clearly $V^0(A, k) = 0$. Define $W^j(A, k) = R(A, k) + \beta \int V^j(A', k(1 - \delta)) Z(dA', A)$ and accordingly S^j, s^j , and a^j , for each $j = 0, 1, \dots, J$. Using Proposition 3, it holds:

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

Lemma 9 For $j \geq 0$, $V^{j+1}(A, \cdot)$ is differentiable with respect to k almost everywhere on the domain $\mathcal{K} \setminus s^j(A)$ for every $A \in \mathcal{A}$

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

$$R_k(A, k^j) - p + \beta(1 - \delta) \int V_k^j(A', (1 - \delta)k^j) Z(dA', A) = 0$$

using the dominated convergence theorem. Then $V^{j+1}(A, k)$ is differentiable in k with its derivative given by: $V_k^{j+1}(A, k) = p$. Plugging in the first-order condition gives Equation 14. If $k > s^j(A)$, $V^{j+1}(A, k) = R(A, k) + \beta(1 - \delta) \int V^j(A', k(1 - \delta)) Z(dA', A)$ and consequently, $V^{j+1}(A, k)$ is differentiable in k with derivative given by Equation 14. \square

Corollary 7 For all $j \geq 0$, $S^j(\cdot)$ and $s^j(\cdot)$ are non-decreasing

Proof For the sequence of function $\{S^j\}_{j \geq 0}$, the proof uses repeatedly Lemma 8. Consider the case $j = 0$. Since $R_{kA} \geq 0$ and $R_k(A, s^0(A)) > p$, $(s^0)'(A) = \frac{R_A(A, S^0(A)) - R_A(A, s^0(A))}{R_k(A, s^0(A)) - p} \geq 0$. Now consider $j \geq 0$. If $s^j(A) = S^j(A)$, the result is immediate so we only consider $s^j(A) < S^j(A)$. Using [8], $V(\cdot, k)$ is differentiable with respect to A almost everywhere for every $k \in \mathcal{K}$. Differentiating $W^j(A, s^j(A)) - W^j(A, S^j(A)) + F + pS^j(A) - ps^j(A) = 0$, it holds:

$$(s^j)'(A) = \frac{W_A^j(A, S^j(A)) - W_A^j(A, s^j(A))}{W_k^j(A, s^j(A)) - p} \text{ if } W_k^j(A, s^j(A)) \neq p$$

The numerator is non-negative from the supermodularity of W^j . $W^j(A, k) - pk$ being differentiable with respect to k almost everywhere, we use the definition of F -concavity for differentiable function: $W^j(A, S(A)) - pS(A) - F \leq W^j(A, s^j(A)) - ps^j(A) + (S^j(A) - s^j(A))(W_k^j(A, s^j(A)) - p)$ which is equivalent to $W_k^j(A, s^j(A)) \geq p$. If the inequality is strict, $(s^j)'(A) \geq 0$. Otherwise, $(s^j)'(A)$ goes to infinity. \square

Proposition 7 $V(A, \cdot)$ is differentiable with respect to k almost everywhere on the domain $\mathcal{K} \setminus s(A)$ for every $A \in \mathcal{A}$

Proof For every $n \in \mathbb{N}$, a sequence of derivatives can be defined recursively as $W_k^{n+1}(A, k) = D(W_k^n)(A, k)$ where the operator D satisfies:

$$D(W_k^n)(A, k) = R_k(A, k) + \beta(1 - \delta) \cdot \left[\int_{\underline{A}}^{a^n(k(1-\delta))} W_k^n(A', k(1 - \delta)) Z(dA', A) + pZ([a^n(k(1 - \delta)), \bar{A}], A) \right]$$

and,

$$\begin{aligned} \|D(W_k^1) - D(W_k^0)\| &= \left\| \int_{\underline{A}}^{a^1(k(1-\delta))} [W_k^1(A', k(1 - \delta)) - W_k^0(A', k(1 - \delta))] Z(A', A) \right. \\ &\quad \left. + \int_{a^0(k(1-\delta))}^{a^1(k(1-\delta))} \{W_k^0(A', k(1 - \delta)) - p\} Z(A', A) \right\| \cdot \beta(1 - \delta) \\ &\leq \beta(1 - \delta) (\|W_k^1 - W_k^0\| + \|W_k^0 - p\|) \end{aligned}$$

Proceeding inductively, for every $n \in \mathbb{N}$, $\|W_k^{n+1} - W_k^n\| \leq \beta(1-\delta) (\|W_k^n - W_k^{n-1}\| + \|W_k^{n-1} - p\|)$. Solving Recursively, for every $n \in \mathbb{N}$,

$$\|W_k^{n+1} - W_k^n\| \leq \beta^n(1-\delta)^n \|W_k^1 - W_k^0\| + \sum_{i=0}^{n-1} \|W_k^{n-i-1} - p\| \beta^i(1-\delta)^i$$

It is well known since [33] that the sequence of functions $\{S^j\}_{j=0}^\infty$ converges uniformly to S . Since the sequence of $\{W^j(A, k)\}_{j=0}^\infty$ converges to W and W is continuous, the sequence of functions $\{s^j\}_{j=0}^\infty$ and $\{a^j\}_{j=0}^\infty$ converges uniformly, respectively, to s and a . Given $\eta > 0$, it exists N_η such that $\|a^{n+1} - a^n\| \leq \eta$ for every integer $n \geq N_\eta$. Hence, for every integer $n \geq N_\eta$:

$$\int_{a^n(k(1-\delta))}^{a^{n+1}(k(1-\delta))} [W_k^n(A', k(1-\delta)) - p] Z(A', A) \leq \|W_k^n - p\| \cdot \|Z([a^n(k(1-\delta)), a^{n+1}(k(1-\delta))], A)\|$$

which implies:

$$\begin{aligned} \left\| W_k^{N_\eta+2} - W_k^{N_\eta+1} \right\| &\leq \beta(1-\delta) \|W_k^{N_\eta+1} - W_k^{N_\eta}\| \\ &\quad + \beta(1-\delta) \|W_k^{N_\eta} - p\| \|Z([a^{N_\eta}(k(1-\delta)), a^{N_\eta+1}(k(1-\delta))], A)\| \end{aligned}$$

Define

$$\begin{aligned} T_\eta &= (\beta(1-\delta))^{N_\eta} \|W_k^1 - W_k^0\| + M_{N_\eta} \\ g(n) &= \beta(1-\delta) \|W_k^n - p\| \|Z([a^n(k(1-\delta)), a^{n+1}(k(1-\delta))], A)\| \end{aligned}$$

It holds: $\|W_k^{N_\eta+2} - W_k^{N_\eta+1}\| \leq \beta(1-\delta) [T_\eta + g(N_\eta)]$. By induction, for every $j \in \mathbb{N}$,

$$\|W_k^{N_\eta+j+1} - W_k^{N_\eta+j}\| \leq \beta^j(1-\delta)^j T_\eta + \sum_{i=1}^{j-1} \beta^i(1-\delta)^i g(N_\eta + j - i)$$

Given $\varepsilon > 0$, there exist N_ε, η and N_η such that $\varepsilon > \beta^j(1-\delta)^j T_\eta + \sum_{i=1}^j \beta^i(1-\delta)^i g(N_\eta + j - i)$ for all $j \geq N_\varepsilon - N_\eta$. Hence, the sequence $\{W_k^j(A_0, k)\}_{j=0}^\infty$ converges uniformly. From the point-wise convergence of $\{W^j(A_0, k)\}_{j=0}^\infty$, the result follows. \square

D The Marginal value of capital

$$\begin{aligned}
& W_k(A_0, k) \\
= & R_k(A_0, k) + \beta(1 - \delta) \left[\int_{\underline{A}}^{a(k(1-\delta))} R_k(A_1, k(1 - \delta)) + \int_{a(k(1-\delta))}^{\bar{A}} p \right] Z(dA_1, A_0) \\
& + a'(k(1 - \delta)) z(a(k(1 - \delta)), A_0) \cdot [R(a(k(1 - \delta)), k(1 - \delta)) \\
& - [W(a(k(1 - \delta)), S(a(k(1 - \delta)))) - pS(a(k(1 - \delta))) + pk(1 - \delta) - F] \\
& + \sum_{j=2}^{\infty} (\beta(1 - \delta))^j \left(\prod_{l=1}^{j-1} Z([\underline{A}, a(k(1 - \delta)^l)], A_{l-1}) \right) \cdot \\
& \left[\int_{\underline{A}}^{a(k(1-\delta)^j)} R_k(A_j, k(1 - \delta)^j) + \int_{a(k(1-\delta)^j)}^{\bar{A}} p \right] Z(dA_j, A_{j-1}) \\
& + a'(k(1 - \delta)^j) z(s^{-1}(k(1 - \delta)^j), A_{j-1}) \cdot [R(a(k(1 - \delta)^j), k(1 - \delta)^j) \\
& - [W(a(k(1 - \delta)^j), S(a(k(1 - \delta)^j))) - pS(a(k(1 - \delta)^j)) + pk(1 - \delta)^j - F]] \\
& + \prod_{l=2}^{j-1} \frac{z(a(k(1 - \delta)^l), A_{l-1})}{Z([\underline{A}, a(k(1 - \delta)^l)], A_{l-1})} \cdot \left(\int_{\underline{A}}^{a(k(1-\delta)^j)} R(A_j, k(1 - \delta)^j) Z(dA_j, A_{j-1}) \right) \\
& + \int_{a(k(1-\delta)^j)}^{\bar{A}} [W(A_j, S(A_j)) - pS(A_j) + pk(1 - \delta)^j - F] Z(dA_j, A_{j-1})
\end{aligned}$$