Essays in Asset Pricing

by

Lorenzo Garlappi

B.A., Università L. Bocconi, Milano, 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES

(Faculty of Commerce and Business Administration, Finance Division)

We accept this thesis as conforming
to the required standard

The University of British Columbia

May 2001

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Commerce

The University of British Columbia
Vancouver, Canada

Date July 27, 2001

DE-6 (2/88)

Abstract

This dissertation consists of two essays dealing with two selected aspects of the investment decision process faced by individuals and corporations.

In the first essay, I develop a model of a multiple-stage patent race between two rival firms to study the impact of technological competition on value and return dynamics of Research and Development (R&D) ventures. The model describes a firm's capital budgeting decision process in the presence of *technical uncertainty*, *market uncertainty* and *preemption*. I characterize the equilibrium of the race and derive optimal investment strategies. Analysis of the equilibrium firm value shows that the premium accruing to the technology "leader" is larger than the loss accruing to the technology "lagger" and that the marginal effect of success/failure is increasing in the uncertainty of cash flows. Risk premia demanded by an ownership claim to competing R&D ventures (i) increase when a rival pulls ahead in the race and (ii) are lower when rivals are "closer" to each other in the development process. Compared to the case where rival firms merge, R&D competition reduces the industry value and lowers the expected completion time for a project. The erosion in value, due to preemption, is higher when firms are "neck-and-neck" and in early stages of development. Numerical simulations show that, in later stages of development, risk premia demanded by the perfectly collusive market are generally lower than risk premia demanded by a portfolio of competing firms. The opposite is true in early stages of development, which suggests that R&D competition may actually lower the cost of early stage financing.

In the second essay, I solve a portfolio allocation problem for an individual who can select between two risky assets and a riskless asset in the presence of capital gains taxes. I treat capital gains taxes as a form of endogenous transaction costs. Using this analogy, I characterize the trading strategy for the two assets, and study the effect of taxes on optimal portfolio diversification. The optimal strategy contains a "no trade" region and a dynamic tax–timing option. I find that the diversification costs due to capital gains taxes are substantial and the value of the tax deferral option is decreasing in the correlation among assets and in the volatility of the risky assets. By comparing the solution of the multiple asset portfolio problem to the one of an investor who can trade only in a mutual fund I am able to measure the value of the flexibility option of the multi-asset case as well as the cost of mutual fund turnover. Finally, I show that imposing a wash-sale constraint generates discontinuous portfolio rebalancing strategies.

Contents

Abstract ii

Contents iii

List of Tables vi

List of Figures x

Acknowledgements xvii

Dedication xviii

1 Introduction 1

1.1 Background for Chapter 2 2

1.1.1 The 'Early Days' 2

1.1.2 A New View of Investments? 3

1.1.3 Strategic Investment Decisions and R&D 5

1.2 Background for Chapter 3 7

1.3 Methodology 8

2 Preemption Risk and the Valuation of R&D Ventures 10

2.1 Introduction 10

2.2 The Patent Race Model 15

2.2.1 Dynamics of the R&D Game 15

2.2.2 Valuation 18

2.3 The Strategic Outcome 19

2.3.1 Analysis of the Model . . . 22
2.4 The Cooperative Outcome . . . 25
2.5 Numerical Results . . . 28
2.5.1 The Effect of Preemption on Value . . . 29
2.5.2 The Effect of Preemption on Risk Premia . . . 34
2.6 Conclusions . . . 42

3 Portfolio Selection with Multiple Assets and Capital Gains Taxes **44**

3.1 Introduction . . . 44
3.2 The Model . . . 48
3.2.1 The Distribution of Asset Returns . . . 49
3.2.2 Investment Strategies . . . 49
3.2.3 Treatment of Capital Gains . . . 51
3.2.4 The evolution of the cost bases and asset holdings over time . . . 52
3.2.5 Feasible Investment Strategies . . . 53
3.2.6 Preferences . . . 56
3.2.7 Dynamic Programming Formulation . . . 56
3.3 Numerical Results for the Two-Asset Case . . . 59
3.3.1 Effect of Taxes on Optimal Diversification . . . 59
3.3.2 Value of Optimal Exercise of the Tax-Timing Option . . . 62
3.4 Effect of Capital Gains Taxes on Fund Trading . . . 65
3.4.1 Capital Gains Taxes and Mutual Funds . . . 66
3.4.2 Tax Cost of Mutual Fund Turnover . . . 70
3.5 Wash Sales Constraints . . . 75
3.5.1 Results . . . 75
3.6 Conclusions . . . 77

Bibliography **79**

Appendices **83**

Appendix A Proofs of Propositions in Chapter 2 **84**

Appendix B Algorithm to solve for $V(n, m, \delta)$ in Chapter 2 97

Appendix C Tables and Figures for Chapter 2 99

Appendix D Feasibility Set for the One-Asset case 113

Appendix E Proofs of Propositions in Chapter 3 116

Appendix F Numerical Implementation for Chapter 3 121

Appendix G Tables and Figures for Chapter 3 124

List of Tables

I Transition Probability Matrix from State (n, m). 18

II The RJV problem 26

III **Parameter Values for the Numerical Solution**. The table lists the values chosen for all parameters required to solve the the patent race model and the RJV model (quarterly values): the number of phases to complete a project (N), the fixed component of the development cost (a), the variable component of the development cost (b), the probability of successfully overcoming an hurdle (π), the average growth (μ) of the potential cash flow process, the volatility of cash flows (σ_δ), the volatility of innovation in the pricing kernel (σ_z), the correlation between innovations in the cash flow process and in the pricing kernel (ρ) and the riskless interest rate (r). The last entry (λ) is the risk premium earned by the completed project and is derived in Proposition 2.3.1. 99

IV **Value functions** $V(n, m, \delta)$, for early stages (Panel A), intermediate stages (Panel B) and late stages (Panel C) of developments of an R&D project requiring 20 stages to be completed. $V(n, m)$ represents both the value of Firm A in subgame (n, m) and the value of Firm B in subgame (m, n). Parameter values are listed in Table III. 105

V **Effect of distance to completion and relative distance between firms on Value of R&D project to firm A.** The relative distance is defined as the gap between the number of stages completed by the two firms $(n-m)$. If $(n-m)>0$ firm A is leader. Panel A considers value functions in the early stages of development (15 phases to complete). Panel B considers value functions in the intermediate stage of development (10 phases to complete) and Panel C considers value functions in the late stages of development (5 phases to complete). Entries marked with '–' represent infeasible relative distances. Parameter values are listed in Table III. 106

VI **Effect of Volatility of cash flows on firm value.** Values are computed in three different stages of the development process, initial stages, intermediate stages and late stages. Panel A considers values for low volatility of cash flows ($\sigma_\delta = 0.10$). Panel B considers values for medium volatility levels ($\sigma_\delta = 0.15$) and Panel C for high volatility level ($\sigma_\delta = 0.30$).Parameter values are listed in Table III. 107

VII **Value-dissipating effect of R&D competition.** The table reports the percentage erosion in value (Φ), computed according to Definition 2.5.1. Panel A contains percentage erosions for early stages of development, Panel B for intermediate stages of development and Panel C for late stages of development. The project requires 20 stage to be completed. Parameter values are listed in Table III. 108

VIII **Impact of Preemption on Investment Decisions.** Panel A reports the value function V and investment strategies when firm A has completed $n = 8$ stages and faces a competitor far behind ("Quasi-monopoly" scenario). It also reports the value function V^J and investment strategy of the RJV. The last column contains the % erosion as defined in Definition 2.5.1. Panel B contains the same information for the case of a "Neck-and-Neck" scenario. Parameter values are listed in Table III. 109

IX **Risk premia $RP(n,m)$ earned on an ownership claim to venture A.** Risk premia are expressed as % p.a. and defined in equation (2.25). Panel A reports risk premia in early stages of development, Panel B reports risk premia in intermediate stages of development and Panel C reports risk premia in late stages of development. Parameter values are listed in Table III. 110

X **Effect of distance to completion and relative distance between firms on Risk Premia of firm A**. The relative distance is defined as the gap between the number of stages completed by the two firms $(n-m)$. If $n-m>0$ firm A is leader. Panel A considers risk premia in the early stages of development (15 phases to complete). Panel B considers risk premia in the intermediate stage of development (10 phases to complete) and Panel C considers risk premia in the late stages of development (5 phases to complete). Entries marked with '–' represent infeasible relative distances. Risk premia are defined in equation (2.25). Risk premia are set equal to zero when the value function is zero.Parameter values are listed in Table III. 111

XI **Risk Premia (% p.a.) for a value-weighted portfolio of R&D firms and for a RJV**. Risk premia RP^{pf} are derived in Definition 2.5.5 and RP^J are derived in Definition 2.5.6. Panel A reports the relationship between RP^{pf} and RP^J in a "Quasi-monopoly" scenario for different stages of development (early, intermediate and late development). Panel B does the same for the case of "Neck-and-Neck"scenarios. The projects requires 20 stages to be completed. Parameter values are listed in Table III. 112

XII **Benchmark Case Parameter Values for the Numerical Solution of the Two-Asset Case**. The investment horizon is T. The parameters μ_i, d_i, and σ_i give the cum-dividend mean, dividend yield, and standard deviation of asset i returns for $i=1,2$. The correlation coefficient is ρ. The tax rate on capital gains and dividends are τ_g and τ_d respectively. The risk aversion parameter is γ, and the subjective discount factor is β. 124

XIII **Value of Tax-Deferral Option**. The table reports the value of the tax-timing option for different levels of the volatility of the two risky assets and for different level of risk aversion. The value is computed as the extra percentage wealth necessary to make an individual paying taxes on accrual indifferent to holding the tax-on-realization portfolio. Fro every case, the option value is computed at the point in which the investor holds the Merton's portfolio and the assets do not embed any gain or loss $(\theta_1=1, \theta_2=1)$. Other parameters are as in Table XII. 125

XIV **Value of Tax-Deferral Option**. The table reports the value of the tax-timing option for different levels of correlation between the two risky assets. The value is computed as the extra percentage wealth necessary to make an individual paying taxes on accrual indifferent to holding the tax-on-realization portfolio. Asset holdings are set such that 20% of wealth is held in each asset (i.e. the Merton portfolio, $x_1 = .2, x_2 = .2$) and the basis to price ratio is $\theta_1 = 1, \theta_2 = 1$. Other parameters are as in Table XII. 125

List of Figures

C.1 **Graphical representation of the patent race**. The tree represents the dynamics of the patent race in subgame $(n, m, \delta(t))$. A and B indicate the two players. In every subgame (n, m, δ), each player simultaneously decides whether to *Invest* (I) or *wait* (W). The decision nodes labeled with a square represent *technical uncertainty.* π^A (π^B) is the probability that firm A (B) is successful when investing. The decision nodes labeled with a circle represent *systematic uncertainty,* i.e., the evolution of the potential cash flow upon completion (δ). The dashed box indicates that the resolution of technical and systematic uncertainty is simultaneous in every subgame. The dashed arcs indicate that $\delta(t + \Delta)$ is drawn from a continuous (lognormal) distribution, as defined in (3.4). Δ is the time interval between two subsequent subgames. 100

C.2 **Family of value functions $V(\cdot, m, \cdot)$ for firm A when B has completed $m =$ 8 stages**. Values of firm A are plotted against δ, the underlying potential cash flow from the completed project and n, the number of stages completed by firm A. Parameter values are listed in Table III. 101

C.3 **Discontinuity in the Value Function**. Panel A shows the value function (as a function of δ) of firm A in subgame (8,6) and Panel B reports the investment strategies of A and B as a function of δ. In Panel B the solid line indicates the investment strategy fro firm A (leader) and the dotted line is the investment strategy for firm B (follower). '0' indicates that a project is mothballed and '1' indicates that a project is activated. Parameter values are listed in Table III. 102

C.4 **Risk Premia of Firm A when firm B has completed $m = 1$ stages.** The risk premia are expressed as a % p.a. and plotted as a function of δ and of the number of stages (n) completed by firm A. The overall project requires 20 stages to be completed. The risk premium is defined as the difference between the continuously compounded expected return of the project and the riskless rate (see equation (2.25)). Parameter values are listed in Table III. 103

C.5 **Risk Premia of Firm A when firm B has completed $m = 8$ stages.** The risk premia are expressed as a % p.a. and plotted as a function of δ and of the number of stages (n) completed by firm A. The overall project requires 20 stages to be completed. The risk premium is defined as the difference between the continuously compounded expected return of the project and the riskless rate (see equation (2.25)). Parameter values are listed in Table III. 104

D.1 **Feasibility Set at time 0.** This figures displays the feasibility set at time $t = 0$ for a one-asset portfolio problem with capital gains taxes. Panel A refers to the case of an embedded capital gain while Panel B refers to the case of an embedded capital loss $B(t) > X(t)$. The shaded region in Panel A is the feasibility set described by equations (D.1) and (D.3). Line a is the self financing constraint (D.1) when wash sales are allowed while the dashed line c is the self financing constraint (D.3) when wash sales are not allowed. Line b is the non-negativity of liquidating wealth requirement (D.3). 114

G.1 **Trades in Stock 1.** This figure shows trades in asset 1 as a function of holdings in asset 1 for a variety of values of the other state variables. Each panel reports the trades for two different levels of the basis to price ratio in asset 1 ($\theta_1 = 0.2, 1.1$) as well as the trades in the absence of taxes. Panel A reports a situation in which asset 2 has a low holding and an embedded gain. Panel B reports a situation in which asset 2 has a low holding and an embedded loss. Panel C reports a situation in which asset 2 has a high holding and an embedded gain. Panel D reports a situation in which asset 2 has a high holding and an embedded loss. Parameter values are as specified in Table XII. 126

G.2 **Regions of No-Trade.** This Figure shows the investment strategy at time $t = 0$ in the two asset problem for a variety of different basis to price ratios. Three different types of actions are detailed as a function of the initial portfolio positions in each asset. In the black region, the investor chooses not to trade in either asset. In the grey region, one asset is traded (bought or sold), while holdings of the other asset remain constant. The white region shows values of the initial holdings where the investor chooses to trade in both assets. The black cross indicates the investment by the non-taxed (Merton) investor. 127

G.3 **Cross-effects of holdings on trades.** This figure shows the trades in asset 1 (α_1) and asset 2 (α_2) as a function of the holding of asset 1 (x_1). For both panels, the holdings of asset 2 are fixed at $x_2 = 0.1$, the basis to price ratio for asset 1 is $\theta_1 = 0.2$ and the basis to price ratio for asset 2 is $\theta_2 = 1.0$. Panel A refers to the case where the two assets are negatively correlated ($\rho = -0.5$), while Panel B refers to the case where the two assets are nearly perfect substitutes ($\rho = 0.9$). The other parameter values are as specified in Table XII. 128

G.4 **Value of Tax-Timing Options.** The value of the tax-timing option is computed as the extra percentage wealth necessary to make the no-tax investor indifferent to holding the taxable portfolio. The figure plots values as a function of the basis and initial holding in asset 1. Every panel plots the option value as a function of holdings in asset 1 (x_1) and its relative basis to price ratio (θ_1). Panel A shows the case with low holdings and a capital gain in asset 2. Panel B shows the case with low holdings and a capital loss in asset 2. Panel C shows the case with high holdings and a capital gain in asset 2. Panel D shows the case with high holdings and a capital loss in asset 2. Parameter values are as specified in Table XII. 129

G.5 **Value of multi-asset flexibility.** This Figure shows the certainty equivalent cost of investing in a mutual fund which which passes all gains and losses to the shareholder. Certainty equivalent costs are defined as the percentage additional wealth necessary to make the mutual fund investor indifferent to holding the two taxable assets individually. Positive values indicate states where the two asset case is preferred. Every panel plots the certainty equivalent costs as a function of holdings in asset 1 (x_1) and its relative basis to price ratio (θ_1). Panel A shows the case with low holdings and a capital gain in asset 2. Panel B shows the case with low holdings and a capital loss in asset 2. Panel C shows the case with high holdings and a capital gain in asset 2. Panel D shows the case with high holdings and a capital loss in asset 2. Parameter values are as specified in Table XII. 130

G.6 **Trades in the Mutual Fund Units.** This figure shows trades in the mutual fund (α_m) as a function of holdings in asset 1 for a variety of values of the other state variables. Each panel reports the trades for two different levels of the basis to price ratio in asset 1 ($\theta_1 = 0.2, 1.5$). Panel A reports a situation in which asset 2 has a low holding and has an embedded gain. Panel B reports a situation in which asset 2 has a low holding and an embedded loss. Panel C reports a situation in which asset 2 has a high holding and an embedded gain. Panel D reports a situation in which asset 2 has a high holding and an embedded loss. Parameter values are as specified in Table XII. 131

G.7 **Value of Index Fund Trading.** This Figure shows the certainty equivalent costs of investing in an index fund which which has tax treatment equivalent to a composite asset in contrast to investing in the two assets separately. Certainty equivalent costs are defined as the amount of additional wealth necessary to make the index fund investor indifferent to holding the two taxable assets individually. The certainty equivalent costs are normalized by initial wealth. Positive values indicate states where the two asset case is preferred. Each panel reports the certainty equivalent cost for three different portfolio allocations as the basis to price ratio in asset 2 (θ_2) varies. Line a refers to a portfolio overbalanced in asset 2. Line b refers to a portfolio equally weighted in the two assets and line c refers to a portfolio overbalanced in asset 1. Panel A shows the case of an embedded gain in asset 1 and Panel B shows the case of an embedded loss in asset 1. Parameter values are as specified in Table XII. 132

G.8 **Tax Costs of Mutual Fund Turnover.** This Figure shows the certainty equivalent cost of investing in an index fund which which has tax treatment equivalent to a composite asset in contrast to investing in a mutual fund which passes all realized gains and losses. Certainty equivalent costs are defined as the amount of additional wealth necessary to make the index fund investor indifferent to holding the mutual fund. The certainty equivalent cost are normalized by initial wealth. Positive values indicate states where the mutual fund case is preferred. Each panel reports the certainty equivalent cost for three different portfolio allocations as the basis to price ratio in asset 2 (θ_2) varies. Line a refers to a portfolio overbalanced in asset 2. Line b refers to a portfolio equally weighted in the two assets and line c refers to a portfolio overbalanced in asset 1. Panel A shows the case of an embedded gain in asset 1 and Panel B shows the case of an embedded loss in asset 1. Parameter values are as specified in Table XII. 133

G.9 **Trades in Stock 1 Under the Wash Sales Constraint.** This figure shows trades in asset 1 as a function of holdings in asset 1 when wash sales are not allowed and for a variety of values of the other state variables. Each panel reports the trades for two different levels of the basis to price ratio in asset 1 ($\theta_1 = 0.2, 1.5$) as well as the trades in the absence of taxes. Panel A reports a situation in which asset 2 has a low holding and an embedded gain. Panel B reports a situation in which asset 2 has a low holding and an embedded loss. Panel C reports a situation in which asset 2 has a high holding and an embedded gain. Panel D reports a situation in which asset 2 has a high holding and an embedded loss. Parameter values are as specified in Table XII. 134

G.10 **Value of Tax-timing option when wash sales are not allowed.** This Figure shows the certainty equivalent cost of investing in two risky assets in the presence of capital gains taxes compared to the no-tax case. Wash sales are not allowed. Certainty equivalent costs are defined as the amount of additional wealth necessary to make the non-taxable investor indifferent to holding assets subject to taxation. Positive values indicate states where the two asset case is preferred. The figure is drawn for the case of an initial holding of asset 2 equal to $x_2 = 0.7$ and for a basis to price ratio $\theta_2 = 1.5$. Parameter values are as specified in Table XII with the exception of the horizon which is set to T=2. 135

G.11 **Value of multi-asset flexibility when wash sales are not allowed.** This Figure shows the certainty equivalent cost of investing in a mutual fund which which passes all gains and losses to the shareholder. Wash sales are not allowed. Certainty equivalent costs are defined as the amount of additional wealth necessary to make the mutual fund investor indifferent to holding the two taxable assets individually. Positive values indicate states where the two asset case is preferred.. Every panel plots the certainty equivalent costs as a function of holdings in asset 1 (x_1) and its relative basis to price ratio (θ_1). Panel A shows the case with low holdings and a capital gain in asset 2. Panel B shows the case with low holdings and a capital loss in asset 2. Panel C shows the case with high holdings and a capital gain in asset 2. Panel D shows the case with high holdings and a capital loss in asset 2. Parameter values are as specified in Table XII. 136

Acknowledgements

This work could not have been completed without the help of many people who, in different ways, assisted me during the years spent in the Ph.D Program at The University of British Columbia. I would like to thank first the members of my thesis committee: Alan Kraus, Margaret Slade and, in particular, my supervisors Vasant Naik and Tan Wang whose outstanding mentorship and guidance were absolutely crucial in the completion of this dissertation. Among the faculty members who were particularly close to me, I would like to express my gratitude to Ron Giammarino and Raman Uppal who always supported my efforts during these years. I would also like to thank some of my fellow doctoral students and friends for their priceless help: Ajay Agrawal, Geneviève Bassellier, Giuseppe Beretta, Luca Bianchi, Cindy Chung, Pietro Ferrario, Lisa Kramer, Yinghu Liu, Martin Martens, John McCarthy, David and Yasmine Moses, Umberto Nebuloni, Dave Nordquist, Marco e Margherita Pattano, Dave Peterson, Olaf Rieck, Amir Rubin, Jacob Sagi, Rushen Shi, Bill and Annerose Sims, Daniel Smith, Sanjeev Swami, Giacinto Tunesi, Don Wagner and, in particular, Joshua Slive whose endless patience in listening to my computer-related questions and research concerns was invaluable. Chapter 3 of my dissertation is part of an ongoing research project on which I am working together with Vasant Naik and Joshua Slive. I would like to thank them for providing a stimulating and enriching intellectual experience.

Finally, I wish to express my immense gratitude to my beloved, Naoko Kato and to my parents.

LORENZO GARLAPPI

The University of British Columbia
July 2001

To my Parents.

Chapter 1

Introduction

Investment decisions are among the most pervasive activities in the life of both individuals and organizations. Firms and government institutions are everyday confronted with investment decisions that can rank from the more ordinary ones (maintenance of the inventory) to the more complex and uncertain (research and development, design and implementation of a welfare program). Similarly, individuals routinely face consumption, saving and investment decisions over their entire lifetime.

In this dissertation I will analyze two selected problems associated with intertemporal investment decisions at the corporate and at the individual level. At the corporate level, I address, in Chapter 2, the issue of evaluating real investment opportunities in the presence of strategic interactions between decision makers. The study is particularly suited to the case of start-up firms competing in Research and Development (R&D). At the individual level, I analyze, in Chapter 3, the portfolio problem of an investor who has to allocate wealth in a multiplicity of risky asset available in the market and is subject to taxes on the realization of capital gains embedded in the traded securities.

The two problems are only apparently unrelated. The solution approach, in fact, relies on recognizing and correctly evaluating the *option-like* features underlying both investment problems. As we will discuss in detail later, the techniques for evaluating corporate real investment opportunities greatly benefited from the awareness of the fact that such opportunities embed a variety of flexibility features that make them similar to more traditional financial options. Similarly, the problem of investing in financial assets in the presence of capital gains taxes contains a great degree of discretionality originated by the fact that taxes are to be paid upon realization and therefore give the tax-payer the *option* to decide when it is best to realize or defer capital gains taxes.

In what follows, I will provide a historical and methodological background for the two problem analyzed in this dissertation. A detailed literature review is contained at the beginning of each chapter.

1.1 Background for Chapter 2

Real investments are defined by the economic doctrine as the "flow of expenditures devoted to increasing or maintaining the real capital stock."[1] At a more micro level, this translates into the activity of contracting an immediate expense in the anticipation of future rewards. For firms, which is where I will focus my attention in Chapter 2, investment decisions are at the heart of the overall resource allocation problem as well as a necessary condition for the existence of the firm itself: firms create value by making good real investment decisions.

1.1.1 The 'Early Days'

The problem of correctly evaluating investments has always attracted the attention of different fields of research. In finance, the study of capital budgeting decisions, i.e. the evaluation of investment projects and consequent allocation of resources to alternative projects, led to one of the most consolidated paradigm for investment valuation: the Net Present Value (NPV) rule. According to this principle, whose origins date back to the seminal work of Irving Fisher (1906, 1930), an investment decision should be undertaken whenever its discounted (expected) cash flow stream is larger than its immediate cost; the difference between the discounted cash flows (DCF) and the investment cost defines the NPV of a project. By undertaking a project with positive NPV, managers are implicitly maximizing shareholders' wealth. This basic tenet, which is a version of the neoclassical paradigm of operating at the point in which the marginal revenue of an investment equate the marginal cost of capital, guarantees the compatibility of the investment decision rule with the existence of a diffuse ownership of a firm and, consequently, can be considered fundamental for the successful operation of a capitalistic economy. The sound economic foundations of the NPV rule made it outlive, at least on pure theoretical grounds, the "competition" of alternative criteria

[1] From *The MIT Dictionary of Modern Economics (1996)*, Fourth edition. Ed. D.W. Pearce.

such as the Internal Rate of Return (IRR) or the pay-back rule, even if, among practitioners, the use of these last criteria is not uncommon.

The study of the relationship between risk and return of financial assets and the consequent derivation of equilibrium asset pricing models (CAPM in particular) provided the NPV paradigm with a proper way of accounting for risk in the valuation of investment projects. A useful way of looking at a firm undertaking a positive NPV project is to consider it as offering its well-diversified shareholders an arbitrage opportunity: a positive NPV means, in fact that the firm is paying less than the expected, and properly risk-adjusted, discounted stream of cash flows. Similarly, the no-arbitrage paradigm for financial securities can be reformulated by saying that financial securities, if correctly priced, are nothing else than *zero* NPV projects.

1.1.2 A New View of Investments?

Notwithstanding its usefulness in capital budgeting decisions and its solid economic foundations, it soon became clear that the Discounted Cash Flow approach, in its original form, could be inappropriate to analyze how capital allocations are linked to long-term firm strategy. In other words, the way in which DCF were calculated implicitly assumed a "passive" management of the investment project during its life. If this may be true in evaluating a bond or a stock (this is where the DCF approach originated), it is definitely a rough, to say the least, approximation for real investment projects. The management has infinite possibilities to alter the evolution of a stream of cash flows from a project during its operation: suspension of the project, abandonment, change of scale and deferment of its initiation are only some examples. This "flexibility" of the management has a value and, as such, it has to be correctly taken into account in any capital budgeting decision. Probably the easiest to see among this flexibility opportunities is the possibility of deferring an investment project. Unlike what is implicitly assumed in the plain NPV rule, a real investment decision is rarely a "now-or-never" decision: there is flexibility about the timing in which the investment can be undertaken. This flexibility is particularly valuable in an uncertain environment in which waiting (and still keeping the opportunity to invest) can be useful to gain more insights on the evolution of uncertainty. This trade-off "invest now–wait" is the same trade-off faced by an investor holding a long position on an American option on a dividend-paying stock (which, in the case or real investment, is represented by the present value of the completed project).[2] This simple intuition, together with the availability of result on derivative pricing since Black and Scholes (1973) and Merton (1973) literately revolutionized the way to look at capital budgeting decisions. It was suddenly clear that many other features of an investment project can be interpreted as options (abandonment, suspensions, sequential decision, change of scale) and priced along the line of the florid literature on financial derivatives. This new view of capital budgeting is now know and referred to by academicians and practitioners as the "Real Options Approach" to investment decisions.

It is important to note, however, that this new view of investments does not contradict the original NPV paradigm. The real options approach is nothing more than a more comprehensive way of computing NPVs. The value of an option, real or financial, is basically obtained via an NPV calculation (the equivalent-martingale-measure or risk-neutral approach in asset pricing is a stark example of this): in financial markets, absence of arbitrage imposes this NPV to be zero. In other words, it is always possible (although maybe burdensome) to encompass within the NPV paradigm any of the valuation procedures that claimed to be based on an "alternative approach". To illustrate this, let us go back to the flexibility option of delaying an investment decision. Every investment project which is not "now-or-never" competes with itself delayed, i.e. it contains a series of "mutually exclusive projects". Investing today is a project, wait and invest tomorrow is another one. Undertaking the latter when the former has a positive NPV does not mean that we are violating the NPV paradigm as long as the latter project has a higher NPV than the former: we are simply undertaking the investment that *maximizes* the NPV. Clarified this, to conform with the existing literature, I will still use the term "Real Options Approach" to refer to an economically correct and forward-looking dynamic NPV approach.

The literature on option valuation of real investments under uncertainty has evolved at an increasing pace in the last two decades. Excellent surveys of this literature are the books by Dixit and Pindyck (1994) and Trigeorgis (1996). Here I remember only some of the path-breaking

[2] It is well-known, see Merton (1973), that the value of an American option on a non-dividend paying stock corresponds to the value of a European option. This means that the exercise strategy for an American option on non-dividend paying stocks is simply to wait until the expiration of the right. The introduction of dividends generates the trade-off above mentioned and makes the timing problem of exercising the option non trivial. Moreover, closed form solutions for the price of American options are available only if the options are *perpetual.*

work in this area. Myers (1977) is the first to introduce the concept of "real options" in the context of growth possibility of a firm and to analyze their interactions with corporate borrowing; Brennan and Schwartz (1985) use an option-based model to value the option to open and close a mine; McDonald and Siegel (1986) provide the first stylized model to study the option to defer investments in continuous-time.

1.1.3 Strategic Investment Decisions and R&D

One of the implicit assumption of the above literature is to view the capital budgeting process as a 'game against nature'. If this assumption is acceptable for firms operating in industries with higher barrier to entry, no close substitutes or unique know-how (e.g. resource extraction, paper products, real estates), it becomes untenable in high-tech industries (e.g. computers, consumer electronics, pharmaceutical) where competition in Research and Development (R&D) plays a major role. In this context, then, most capital budgeting problems will involve almost inevitably also a strategic game against competitors. The original NPV approach and its real options extension have been looking at investment opportunities as monopoly rights: this justifies the focus of that literature on the optimal exercise strategy of proprietary investment opportunities. To better understand investment decisions in industries characterized by a high degree of competition, though, we cannot look only at how investment options are exercised but also at (i) how these options are acquired and (ii) how their exercise is affected by the presence of competitors holding similar rights.

It is not hard to think of real world situations of investments in which the payoff is awarded to the first to undertake an investment option (*patent races* and *takeovers* are only some examples). In this framework, the valuation problem acquires a strategic dimension that goes beyond the traditional results available from the valuation of financial options and calls for a more complex game-theoretic approach.

Within the class of strategic investment decisions, R&D investments represent a particularly interesting problem. The valuation of these investment is probably one of the most challenging problems in financial economics. This complexity arises from the high degree of uncertainty generated by the fact that R&D investments produce cash flows only after potentially many (and unpredictable) stages of research. Moreover, every single stage cannot be looked at as a stand-alone investment but as a prerequisite for potential completion of future stages (sequential investments or compound options). The problem becomes even more involved if we take into account the possibility of "races" in the investment game in which two or more firms compete to bring an innovation to the market. Examples of this kind of competition are everyday news in high-tech industries. To name only some we can think of the development and commercialization of faster computer processors which involved Transmeta and Intel[3] and of the competition between SkyBridge and Teledisc on the development of high speed data transmission satellites.[4] A very recent stark example of a patent race is represented by the competition between Celera Genomics and the Human Genome Project on the decoding of the human genome.[5] The importance of R&D for macroeconomic growth, the increasing presence of R&D companies on organized exchanges and the existence of competition in R&D make the problem of correctly analyzing and evaluating strategic capital budgeting decision in R&D a compelling challenge.

The purpose of Chapter 2 is to develop a dynamic game-theoretic model to study a firm's capital budgeting decision by explicitly accounting for the presence of rivals in the investment process. I borrow from the economic literature the techniques to model a patent race problem and from the real option literature in finance the intuition of looking at a sequential investment as a compound option. This will allow me to to derive implication of R&D competition on the valuation and expected returns (cost of capital) of firms engaged in a technology race. This approach enriches the traditional analysis of a patent race by adding an additional component of uncertainty, the market, or systematic, uncertainty, which transforms the definition of the "prize" in a race from a static object into a "moving target". Secondly, my model allows me to analyze how different sources of uncertainty (systematic and idiosyncratic) interact in the determination of returns and risk premia of R&D ventures. Finally, by investigating the problem of a perfectly coordinate research program, as opposed to a technology race, I can assess the impact of "preemption risk" on the value of competing firms as well as on their cost of capital.

Does competition generate risk-premia in stocks of companies that engage in investment races? How does the stock of a firm engaged in a race react to the announcement of the successful completion of a phase in a sequential investment project? Is investing in a coordinate research project less or more risky than investing in a portfolio of competing ventures?

[3]See *New York Times*, January 20, 2000 and October 11, 2000.

[4]See *Business Week*, August 30, 1999 and *New York Times*, June 4, 2000.

[5]See *New York Times*, September 15, 1998 and *Business Week*, June 12, 2000.

This questions will be addressed in Chapter 2 in the context of a non-cooperative sequential investment game of R&D with technical *and* market uncertainty.

1.2 Background for Chapter 3

The study of investment decisions faced by households represents one of the classical problems in financial economics. The problem for an individual is to find the optimal allocation of wealth among available investments opportunities. In solving such a problem the household has to decide (i) the amount of wealth to allocate between consumption and savings and (ii) how to allocate the saved wealth in the existing assets. The study of the problem of choosing the optimal investment mix is referred to as *portfolio-selection theory*. There is an extensive body of literature that deal with this problem, starting with the seminal work of von Neumann and Morgenstern (1947) who axiomatize the preferences of an individual using the expected utility criterion to make investment decisions under uncertainty. Given the amount of attention that the portfolio-selection problem has received, an overview of the studies in this direction is beyond the scope of this introduction. Perhaps two of the most influential contributions in this area are represented by the work of Hakansson (1970) and Merton (1971). The former addresses the portfolio selection problem in a discrete-time model while the latter provides a clean formulation and solution in continuous-time under some assumption about preferences (Hyperbolic Absolute Risk Aversion) and return distribution (Geometric Brownian Motion).

After the elegant solution provided by Merton (1971), increasing effort has been devoted to the solution of the problem in the presence of frictions, in particular, transaction costs and short-selling constraints. It is only towards the end of the '80s that the first satisfactory results appear. Costantinides (1986) provides an approximate solution of the portfolio problem under transaction costs while Davis and Norman (1990) and Dumas and Luciano (1991) are the first to provide provide a characterization of the solution of the portfolio problem with a single risky asset and with proportional transaction costs. Akian, Menaldi and Sulem (1996) extend Davis and Norman's (1990) results to the multi-asset case.

In Chapter 3, I rely on ideas from the transaction costs literature to address the intertemporal portfolio allocation problem in the presence of capital gains taxes. At first sight, taxes can be thought of as a transaction cost upon selling. At a deeper level, however, it is evident that these transaction costs have a highly complex nature. Firstly, selling an asset at a loss can actually generate transaction revenue in the form of negative taxation. Secondly, the magnitude of the cost or benefit is itself dependent on the state variables of the problem (i.e. the portfolio holdings and the the tax bases of the assets). Finally, given that the investor can decide not to incur in capital gain taxes, the burden of such frictions is at the discretion of the household. This last aspect gives the investor a valuable tax-timing option which will be traded off with optimal diversification in the solution of the optimal portfolio allocation problem.

What is the optimal tax-realization strategy when trading in multiple assets in the presence of capital gains taxes? How much are diversification benefits affected by capital gains taxes? How can we assess the value of the tax-timing and flexibility option? In Chapter 3, I address the above questions by solving the problem of an investor who maximizes his or her lifetime utility of wealth and can trade in multiple assets in the presence of capital gains taxes. The analysis provides a characterization of the optimal trading strategy that bears some similarities with the strategy emerging in a portfolio problems with transaction costs. The choice of a multiple asset setting is dictated by the desire of analyzing the trade off between optimal tax-realization and optimal diversification. This trade off is investigated in greater detail by considering the portfolio problem of investors with different level of sophistication. In particular I focus on the investment problem for a household facing the opportunity of investing in a mutual fund and in an index fund in the presence of capital gains taxes. The tax treatment of those instruments allows me to decompose the driving forces behind the trade-off between optimal tax realization and optimal diversification and to assess the magnitude of the tax-timing option embedded in the portfolio problem.

1.3 Methodology

Conceptually, the two problems analyzed in this dissertation share similar characteristics, i.e., as discussed earlier, the presence of option-like features that need to be correctly recognized and evaluated. In Chapter 2 such options are "real options" characterizing the investment opportunity of a R&D firm and whose value is affected by uncertainty in the market as well as by the presence of strategic competitors. In Chapter 3 such options are embedded in the opportunity available to the household to defer the realization of capital gains.

I analyze both problem in a discrete-time intertemporal setting. To address the valuation problem in this context, I rely in both chapters on Dynamic Programming techniques. In Chapter 2 I extend the value iteration algorithm (see Bellman (1957) or Puterman (1994)) for infinite-horizon problems to account for the presence of multiple decision makers acting strategically. Chapter 3, on the other side, relies on a more traditional backward induction algorithm for the solution of finite-horizon problems in the presence of uncertainty.

Chapter 2

Preemption Risk and the Valuation of R&D Ventures

2.1 Introduction

The appraisal of Research and Development (R&D) projects is a crucial issue in any modern economy. A recent study by the Directorate for Science, Technology and Industry (DSTI) within the Organization for Economic Cooperation and Development (OECD) shows that in OECD countries the "expenditure on R&D reached almost \$500 billion in 1997, or more than 2.2% of OECD-wide GDP, with a strong increase in spending in the second half of the 1990s. The global research effort is larger than ever before." (See DSTI (2000) and OECD (1999)).[1]

The valuation of R&D investments represents a challenging problem in financial economics. This challenge arises from some of the peculiar features shared by such investments. R&D investments are characterized by multiple, qualitatively different sources of uncertainty and produce cash flows only after potentially many (and unpredictable) *stages* of research. The progress towards completion requires therefore a sequence of successful investments each of which opens the possibility to undertake the next phase of research. The valuation problem is further complicated if we take into account the possibility of *races* within the R&D process in which two or more firms compete

[1]In the United States, data from the National Science Foundation (NSF) show that industry-funded R&D as a percentage of GDP has risen 70 percent during the past two decades reaching the level of \$144 billion (1.7% of GDP) in 1998. As the OECD recognizes, today the "innovative efforts, and R&D in particular, are undoubtedly the major factors behind technical change and long-term economic performance" (see OECD (1998)).

to bring a new product or technology to the market and only the winner can capture the final cash flows generated by the completed project. Examples of R&D competition are everyday news in knowledge-intensive and high-tech industries. Among the most well-known are the development and commercialization of a pharmaceutical drug and of faster computer processors.

The relevance of R&D for macroeconomic growth, the highly competitive environment in which some R&D ventures operate and the increasing presence of stocks of R&D companies on organized exchanges make the problem of correctly analyzing and evaluating *strategic capital budgeting* decisions in R&D a compelling effort for both researchers and practitioners.

In this article, our objective is to analyze and quantify the effect of *strategic preemption* on the value and return dynamics of R&D ventures engaged in a patent race. We address this issue at two levels. First, we model the capital budgeting process of an R&D firm engaged in a technology race as a stochastic game. The solution of this game allows us to define the equilibrium value and risk premia demanded by an ownership claim in such ventures. Second, we solve for the perfectly cooperative case in which investments in R&D are coordinated between research units. The case of cooperation in R&D provides a benchmark that is used to analyze the effects of preemption in technology races on (i) investment decisions, (ii) asset prices and (iii) risk premia.

Our contribution to the existing literature in R&D capital budgeting is on several fronts. First, the capital budgeting process is extended to explicitly take into account strategic exercise of investment options during the patent race. This allows us to analyze how value and risk premia evolve as a consequence of the strategic positioning of firms within the race. This analysis is not possible without modeling the investment as a strategic game. Second, by casting the capital budgeting process in an asset pricing framework we can identify and quantify how and to what extent the different sources of risk embedded in a patent race contribute to the overall riskiness of an R&D venture. Third, by comparing the strategic outcome to the collusive outcome we have an intuitive way of assessing the effect of preemption on value, investment decisions and risk premia in different stages of the development process. Finally, besides developing a tractable model in which the impact of preemption on pricing can be evaluated, we also provide a method for solving infinite-horizon stochastic games and characterize some of the properties of the equilibrium.

Berk, Green and Naik (1999) develop a valuation model for R&D ventures by analyzing the interaction between different sources of risk in the investment process: (i) *Technical uncertainty,* i.e., the uncertainty related to the success of the research process itself, (ii) *Systematic uncertainty*, which refers to industry-wide, macroeconomic conditions that affect the overall profitability of a venture (iii) *Learning by doing*, i.e., the fact that resolution of uncertainty on length and cost of R&D is resolved through investment and (iv) *Competitive threat or obsolescence*, modeled in a reduced form as a probability of a "catastrophic" event (e.g. the entrant of a rival developing a superior product) that terminates the investment process. Following the patent race models of Grossman and Shapiro (1987) and Harris and Vickers (1987), our model builds on Berk, Green and Naik (1999) to explicitly account for strategic interactions in the investment process of firms engaged in a patent race as well as different sources of risk embedded in the investment process. The risk of being preempted by a rival in the development and commercialization of the results of the R&D process will be referred to as *preemption risk.*

Consistent with Berk, Green and Naik (1999) we find that, despite the idiosyncratic nature of technical uncertainty, the distance to completion contributes to the determination of required risk premia. By modeling the investment process as a game, we are also able to identify another crucial idiosyncratic determinant of value and risk premia: *the relative position within a race.* The intuition behind these results relies in the *option-like* nature of the decision to mothball a research project. The resolution of technical uncertainty, as well as the strategic positioning within the race, alter the value and properties of such options. Since the decision to exercise these options is based also on the evolution of the underlying potential profitability of the project (systematic risk), we find that purely idiosyncratic factors *indirectly* contribute to the determination of the overall riskiness of a competing venture. In particular, while the successful overcoming of a R&D stage reduces the "leverage" (i.e. the expected cost to completion) of the investment option of the successful firm, for the preempted firm this is equivalent to an *increase* in the leverage of its investment option. The "increased leverage" effect of a failure is due to the fact that the follower in a race tends to invest at higher level of cash flow. Losing a stage is therefore equivalent to increasing the expected time to completion and, as a consequence the leverage of the investment option.

The comparison of values in the patent race to those in the perfectly collusive case also unveils the well-documented value-dissipating effect of R&D competition, which is more dramatic when rival firms are close to each other in the technology race. However, despite this, risk premia on a portfolio of competing R&D ventures are not substantially different from risk premia earned in the collusive case, with a tendency of the former to be lower in early stages of research. *Early* R&D competition lowers the systematic risk of a portfolio of competing firm.

Our work is related to two major strands of literature in economics and finance. In economics, there is a wide variety of models designed to analyze situations in which firms compete to bring innovations to the market (patent races). Loury (1979), Lee and Wilde (1980) and Reinganum (1981, 1982) focus on the technical uncertainty in the relationship between effort and success, and look at the patent race problem as a stationary game. This approach, which emphasizes the relationship between market structure and level of innovation activity, prevents, on the other side, a direct analysis of the effect of strategic interactions on the investment decision of competing firms. The issue of strategic interactions among players has been pioneered by Fudenberg et al. (1983) and Harris and Vickers (1985), who analyze multiple-staged patent race models with no uncertainty, the former in the context of a simultaneous move game and the latter in a sequential move game. The exclusion of uncertainty, though, leads these authors to conclude that the behavior of a winner is as if he/she is facing no competition at all. In the absence of uncertainty in fact, the winner has a *credible threat* of undoing any effort made by the rivals which deters them from competing. The first attempts made to combine strategic interaction within a game and technical uncertainty are the works of Grossman and Shapiro (1987), who develop a model of dynamic two-stage R&D competition, Judd (1985) who presents a continuous-time model of a race with a continuum of states, and Harris and Vickers (1987) who extend Grossman and Shapiro's (1987) result to a multiple-stage race. As we mentioned above, our model is closer in spirit to Grossman and Shapiro (1987) and Harris and Vickers (1987). As in these models we capture the strategic interactions between players within a non-cooperative game. In addition, by explicitly modeling the evolution of potential cash flows generated upon completion we are able to assess the the impact of preemption on value and quantify the interaction of the different sources of risk in the determination of the overall riskiness of the venture.

The second strand of literature in which our model naturally falls is the *real-option* approach to capital budgeting. This approach, pioneered by Brennan and Schwartz (1985), has recently received considerable attention in the finance literature (the books by Dixit and Pindyck (1994) and Trigeorgis (1996) provide a good survey). One of the implicit assumptions of the real-option approach, though, is to view the capital budgeting process as a "game against nature". While this may be reasonable for firms operating in industries with low competitive pressure, it is less tenable in industries where competition in R&D plays a major role (e.g. computers, consumer electronics, pharmaceutical). In this context, then, most capital budgeting problems will, almost inevitably, involve a strategic game against competitors.

Attempts of studying strategic investment decisions in a real-option framework can be roughly classified into two categories. In the first category, the threat of competition is considered in a *reduced form* and modeled as an exogenous factor altering either the drift of the present value of the completed project (Trigeorgis (1990), Schwartz and Moon (1995), Childs and Triantis (1997), Grenadier and Weiss (1997)), or the stochastic evolution of the underlying cash flows (Berk, Green and Naik (1998)). In the second category, which is closer to our approach, competition is considered in *structural form.* Each investor's decision is contingent upon the others' actions and the option-exercise problem is modeled as a dynamic game. One of the earliest work in this direction is Smets (1991), who develops a sequential investment model in a duopoly with a stochastic dynamic environment. Given that in his model investment opportunities are always available, Smet's framework is not entirely suitable to model a patent race. Smit and Trigeorgis (1993) focus on the characteristics of the equilibrium in the product market that results as a consequence of a preliminary R&D competition. Their major interest is in analyzing the policy effects of proprietary, versus shared benefit, of R&D. Lambrecht and Perraudin (1997) study an option-exercise game in continuous time with the traditional methodology of contingent claim analysis. In the context of a one-stage patent race with incomplete information and no technical uncertainty, they show that firms may find it optimal to win a patent race and let the patent "sleep". Grenadier (1999a) looks at a one-stage incomplete-information option-exercise game to obtain conditions underlying "rational herding" in investment strategies. Finally, Grenadier (1999b) provides a general framework for deriving equilibrium investment strategies in continuous-time within an oligopolistic industry. Our emphasis is on the pricing effect of strategic behavior of competing firms. Unlike Smit and Trigeorgis (1993) we do not model the product market explicitly other than through its contribution to the systematic risk faced by competing firms. Unlike Lambrecht and Perraudin (1997) and Grenadier (1999a,b), we use a discrete-time, complete-information model in which we explicitly account for the different sources of uncertainty embedded in the R&D process. This allows us to develop a tractable asset pricing model that we use to to analyze (and quantify) the impact of preemption on security prices and returns.

The rest of the paper is arranged as follows. Section 2.2 formalizes the model used to describe investment decisions of firms engaged in a patent race. In Section 2.3, we develop a valuation technique for competing R&D ventures and derive analytical properties of the solution to the patent race. Section 2.4 analyzes the case of coordination in the R&D process. This case will be used as a benchmark to study the impact of preemption on value and return dynamics. In Section 2.5 we perform a numerical simulation of the model that allows us to assess the impact of R&D competition on asset prices, investment decisions and risk premia. Section 2.6 concludes. Proofs of the main propositions are collected in Appendix A, while Appendix B contains a description of the algorithm used in our numerical computations.

2.2 The Patent Race Model

In this section we develop a non-cooperative game-theoretic model of R&D capital budgeting decisions. We describe the dynamics of the R&D game and the procedure used to evaluate future random stream of cash flows. In the next Section we formalize the concept of equilibrium and optimal investment strategies.

2.2.1 Dynamics of the R&D Game

Consider two all-equity financed, single-project firms (A and B) competing in the development of a project that requires N phases to be completed. As an example we can think of the process of drug development: the regulation imposed by the Food and Drug Administration (FDA) before approval of a drug, makes this process highly standardized, with multiple stages that are reasonably easy to identify.[2] No prizes are awarded for intermediate success (intermediate patents): the only non-negative cash flow will originate at the end of the race to the winning firm. To focus on the investment problem we look at this project in isolation, as if it were the only asset on the firm's balance sheet. we assume that firms have enough internal resources to undertake investments when

[2]See Schwartz and Moon (1995).

it is optimal to do so.[3]

The first firm to complete the project will gain a perfect patent protection on the product developed and will assume *de facto* a monopoly position in the product market.

A graphical representation of the game is provided in Figure C.1. At each date before completion, the two firms must decide, *simultaneously*, whether to keep working on the project (I for *invest*) in the attempt to reach the next hurdle, or to mothball (W for *wait*). In making their decisions on whether to undertake a phase of the investment or not, the firms consider: (i) their position in the investment race (i.e. the number of stages completed) and (ii) a signal in the form of *potential* cash flows δ generated by the completed project.

As in Berk Green and Naik (1999), we model δ as a geometric random walk

$$\delta(t+\Delta) = \delta(t)\exp\left(\mu\Delta - \frac{1}{2}\sigma_\delta^2\Delta + \sigma_\delta\sqrt{\Delta}\varepsilon_\delta(t+\Delta)\right) \tag{2.1}$$

where Δ denotes the time interval between two subsequent realizations of δ and $\varepsilon_\delta(t+\Delta)$ is a standard normal random variable. The stochastic process δ is defined on the probability space $(\Omega, \mathcal{F}^\delta, P)$ where Ω is the event space, $\mathcal{F}^\delta$ is the filtration generated by (2.1) and P is the probability law of δ derived by (2.1). The signal δ is common knowledge. Even though the firms involved in the race are evidently not generating cash flow before completion we will still refer to δ as to the "cash flow" process. This process can be interpreted as an indicator of general economic conditions for the industry in which the two firms compete. Its role is to capture the *systematic* component associated with the ultimate cash flows. As we will see, this will add a non-trivial option component to the investment decisions in the innovation race.

We assume that when a firm decides to continue with the current research phase, it must bear an investment cost $I(t)$ which contains both a fixed and a variable component:

$$I(t) = a + b\delta(t). \tag{2.2}$$

The investment, or development, cost represents the salary to be paid to research scientists, development engineers as well as the cost of equipment, market research, etc. The variable component of the investment cost reflects both the influence of macroeconomic factors (through δ) such as

[3]It may be useful to think of these competitors as start-up firms. Such firms generally cannot raise financing in the debt market or public equity market. For these kind of firms the *private equity* market (e.g. venture capital) represents most of the time the unique form of financing.

inflation and wage rates and the possible existence of incentive contracts linking employee compensation to forecasted profitability of the project. The fixed component simply captures those portion of the incurred investment cost that are not dependent of δ.

Techical uncertainty is modeled as follows. We assume that, once a decision to invest is undertaken by firm $i = A, B$, the probability that i will be successful in the next Δ instants is $\pi^i\Delta$ with π^i common knowledge.[4] For simplicity, we assume that the probability of overcoming a hurdle is independent of the particular stage of the investment.

Let $n(t)$ be the random variable representing the number of phases completed by firm A at time t and $m(t)$ be the analogous variable for firm B. Every firm knows the number of stages completed by the opponent, i.e $(n(t), m(t))$ is common knowledge at every time t.[5] For tractability, we assume that the probability of having the two firms successfully completing a stage contemporaneously is negligible. In other words, we exclude simultaneous success in the investment race when both players decide to invest. If the interval Δ is "small enough", this is an acceptable approximation. If we think, for example, of the process of developing a new drug or a new computer processor and assume a period Δ of a quarter, it is not unreasonable to exclude the likelihood of having two research labs contemporaneously making major achievements *in the same quarter* towards the completion of their projects. Formally, we may think of the decision to invest as a decision to activate a Poisson process; it is known that, for two independent Poisson processes, the probability of a simultaneous arrival in the interval Δ (i.e. the probability of both players succeeding in the interval Δ) is of order $o(\Delta)$. The "approximate" transition probability matrix for the process $(n(t), m(t))$ is given in Table I, where $[W, W]$, $[W, I]$, $[I, W]$ and $[I, I]$ represent the prevailing investment decisions at time t of the two firms.

If we look at the first two columns of the approximated transition probability matrix, we see that the probability of successfully overcoming a hurdle for a firm who decides to invest is independent of the action of the opponent. This is acceptable provided we interpret the investment

[4]The probability of succeeding is unaffected by the investment decision, if we exclude the fact that it is zero if no investment is made. In other words, we are modeling a linear effort function, with the natural "bang-bang" decisions represented by I and W. This distinguishes our model from many models in the patent race literature (see Grossman and Shapiro (1987) and Harris and Vickers (1987) for examples).

[5]It may be useful to think of $n(t)$ ($m(t)$) as the accumulated knowledge, at time t, by firm A (B). In such an interpretation, common knowledge of $(n(t), m(t))$ means that a firm knows, or can figure out, *how much* the opponent knows, without knowing *what* the opponent really knows. Examples of this situation abound in academics.

Table I. Transition Probability Matrix from State (n, m).

	$(n+1, m)$	$(n, m+1)$	(n, m)
$[W, W]$	0	0	1
$[W, I]$	0	$\pi^B\Delta$	$1 - \pi^B\Delta$
$[I, W]$	$\pi^A\Delta$	0	$1 - \pi^A\Delta$
$[I, I]$	$\pi^A\Delta$	$\pi^B\Delta$	$1 - \pi^A\Delta - \pi^B\Delta$

decisions of the two firms as purely technical research efforts. There is no reason why the success of research for firm A should be affected by research efforts of B. In contrast is the case in which the investment decision of a player has an effect on the potential market shares of the competing firms. We believe that the first interpretation is more suitable for early stage development efforts.

2.2.2 Valuation

To evaluate a stream of random future cash flows, we introduce a pricing kernel in the economy. We will assume that the pricing kernel $z(t)$ evolves according to a geometric random walk:

$$z(t+\Delta) = z(t)\exp\left(-r\Delta - \frac{1}{2}\sigma_z^2\Delta - \sigma_z\sqrt{\Delta}\varepsilon_z(t+\Delta)\right) \tag{2.3}$$

with r denoting the riskless interest rate and $\varepsilon_z(t+\Delta)$ a standard normal random variable. I finally assume that for every t, $cov(\varepsilon_\delta(t), \varepsilon_z(t)) = \rho$. The systematic risk of an uncertain stream of cash flows is measured by taking covariances with the pricing kernel. We do not model the term structure of interest rates explicitly and assume r to be constant.[6]

The existence (and uniqueness) of a pricing kernel implies the existence of a probability measure Q, equivalent to P, such that, for every random variable $\tilde{x}(t+s)$ defined on the probability space $(\Omega, \mathcal{F}^\delta, P)$, $E_t^Q[\tilde{x}(t+s)] = E_t^P[\frac{z(t+s)}{z(t)}\tilde{x}(t+s)]$ for all t and s, i.e., Q is an *equivalent martingale measure.* The following proposition explicitly defines the expected values of future cash flows under the equivalent martingale measure Q.

Proposition 2.2.1 *The value at time t of the future cash flows $\delta(t+s)$, under the equivalent*

[6] By exogenously specifying a pricing kernel, we am ruling out the possibility that firms' investment decisions may alter the span of the economy. Our approach is therefore a partial equilibrium approach. In the context of small, start-up firms engaged in technology races this assumption is justifiable. A thorough investigation of the effect of technological innovation on security prices would require an analysis of incomplete markets in a general equilibrium approach (see Wang (1999)).

martingale measure Q, can be computed as

$$E_t^P\left[\frac{z(t+s)}{z(t)}\delta(t+s)\right] = \delta(t)\exp[s(\mu - r - \lambda)] = \delta(t)\beta^s \tag{2.4}$$

where

$$\beta = e^{\mu-(r+\lambda)} \tag{2.5}$$

is the risk-adjusted discount factor and

$$\lambda = \rho\sigma_\delta\sigma_z \tag{2.6}$$

is the parameter measuring the price of risk.

From the above proposition it is immediate to obtain the value of a completed project (See Proposition 2.3.1). Our focus in the next section will be to derive a valuation technique for an R&D venture facing a rival *before* completion of the race.

2.3 The Strategic Outcome

The game outlined above belongs to the class of discrete-time nonzero-sum stochastic games. The evolution of a stochastic process—in our case the triplet $\{n(t), m(t), \delta(t)\}_t$— determines the conditions under which every round of the game is played. Because, by assumption, the number of stages completed by both firms is common knowledge, we would expect the equilibrium strategy of a firm to depend upon the position of the opponent. The solution of this game involves (i) the derivation of the value function for each player and (ii) the determination of the equilibrium strategies prevailing in each sub-game. In this section, we provide the formal setup for a stationary solution of such a game.

Assume that the state of the game is represented by the triplet (n, m, δ). The action space of firm A in (n, m, δ) is summarized by the indicator variable $u \in \{0, 1\}$ where $u = 1$, if firm A invests (I) and $u = 0$, if firm A mothballs (W). Similarly, firm B's action space will be denoted by $\nu \in \{0, 1\}$.

An *investment strategy* for a firm is a rule that tells the firm what action to choose in every possible state. A *mixed* strategy is a probability distribution over the action space. In a general stochastic game this choice may depend on the history of the game. For my purposes, we consider only *stationary Markov strategies*, i.e., strategies which are invariant with time and depend only on the current state.

Let (f, g) be a pair of Markov strategies. We denote by $V^A_{(f,g)}(n, m, \delta)$ $(V^B_{(f,g)}(n, m, \delta))$ the value to firm A (B) in state (n, m, δ), *given* the strategy (f, g).[7] Since at the end of the race there are no strategic issue involved, the value of the completed project is easily determined from Proposition 2.2.1. Obviously, by the winner-take-all provision, if firm A (B) is the winner of the race, the value to firm B (A) is zero.[8] I summarize this in the following proposition

Proposition 2.3.1 *If firm A completes the project first, then*

$$V^A(N, m, \delta) = \frac{\delta}{1-\beta}, \quad m < N; \tag{2.7}$$

$$V^B(N, m, \delta) = 0, \quad m < N. \tag{2.8}$$

Similarly, if firm B completes first, [9]

$$V^B(n, N, \delta) = \frac{\delta}{1-\beta}, \quad n < N \tag{2.9}$$

$$V^B(n, N, \delta) = 0, \quad n < N. \tag{2.10}$$

Given the strategic nature of the problem, the equilibrium values and investment strategies of the two competing firms have to be determined *simultaneously* in a dynamic optimization framework. The following proposition characterizes the equilibrium in the R&D game.

Proposition 2.3.2 *Let (u_*, ν_*) be an equilibrium strategy pair point for the R&D game. Then the equilibrium (stationary) values $(V^A_{(u_*,\nu_*)}, V^B_{(u_*,\nu_*)})$ of the competing R&D ventures are given by solving the following system of recursions*[10]

$$V^A_{(u_*,\nu_*)}(n, m, \delta) = \max \Big\{ \pi^A E^Q_\delta [V^A_{(u_*,\nu_*)}(n+1, m, \tilde{\delta})] + \nu_* \pi^B E^Q_\delta [V^A_{(u_*,\nu_*)}(n, m+1, \tilde{\delta})]$$

[7]The existence of such functions is guaranteed by the fact that, *given* an investment strategy (f, g), the value of such strategy pair is recursively defined as the expectation, under the equivalent martingale measure, of the value next period. If the discount rate is positive, such recursion is a contraction mapping and admits a unique fixed point.

[8]The winner-take-all provision is a case of *infinite* patent protection. In a separate article I analyze the the impact of the *length* of patent protection, by relaxing the winner-take-all requirement.

[9]*Notice that the condition $V^i(N, N, \delta)$, $i = A, B$ is not specified, since the state (N, N) cannot be reached under the assumed transition probability matrix (see Table I).*

[10]*We indicate with $\tilde{\delta}$ the (random) cash flow in the next period defined in (2.1). We also assume for simplicity $\Delta = 1$. This, of course, imposes that $\pi^A + \pi^B < 1$, to avoid negative probabilities. In a general setting we can always choose Δ small enough to guarantee $1 - \pi^A\Delta - \pi^B\Delta \geq 0$.*

$$+(1-\pi^A-\nu_*\pi^B)E_\delta^Q[V^A_{(u_*,\nu_*)}(n,m,\tilde{\delta})]-a-b\delta,$$
$$\nu_*\pi^B E_\delta^Q[V^A_{(u_*,\nu_*)}(n,m+1,\tilde{\delta})]+(1-\nu_*\pi^B)E_\delta^Q[V^A_{(u_*,\nu_*)}(n,m,\tilde{\delta})]\Big\} \quad (2.11)$$

$$\begin{aligned} V^B_{(u_*,\nu_*)}(n,m,\delta) = \quad & \max\Big\{\pi^B E_\delta^Q[V^B_{(u_*,\nu_*)}(n,m+1,\tilde{\delta})]+u_*\pi^A E_\delta^Q[V^B_{(u_*,\nu_*)}(n+1,m,\tilde{\delta})] \\ & +(1-u_*\pi^A-\pi^B)E_\delta^Q[V^B_{(u_*,\nu_*)}(n,m,\tilde{\delta})]-a-b\delta, \\ & u_*\pi^A E_\delta^Q[V^B_{(u_*,\nu_*)}(n+1,m,\tilde{\delta})]+(1-u_*\pi^A)E_\delta^Q[V^B_{(u_*,\nu_*)}(n,m,\tilde{\delta})]\Big\}, \end{aligned} \quad (2.12)$$

with boundary conditions

$$V^A(N,m,\delta) = \frac{\delta}{1-\beta}, \quad m<N; \quad (2.13)$$
$$V^B(n,N,\delta) = \frac{\delta}{1-\beta}, \quad n<N; \quad (2.14)$$
$$V^A(n,N,\delta) = 0, \quad n<N; \quad (2.15)$$
$$V^B(N,m,\delta) = 0, \quad m<N, \quad (2.16)$$

where

$$E_\delta^Q[V^i_{(u_*,\nu_*)}(n,m,\tilde{\delta})]=E^P\left\{\frac{z(t+1)}{z(t)}V^i_{(u_*,\nu_*)}(n,m,\tilde{\delta})\Big|\,\delta\right\} \quad \textit{for all } (n,m)\in\{0,1,\dots,N\}^2.$$

Equations (2.11) and (2.12) describe the Nash equilibria in the subgame (n,m,δ). The first factor inside the $\max\{\cdot\}$ operator represents the payoff from undertaking the investment in the next stage and the second factor is the payoff from mothballing. Both these payoffs are dependent upon the investment decision of the opponent. Note that valuation is carried out by using the equivalent martingale measure Q. The boundary conditions (2.13)-(2.16) are derived explicitly in Proposition 2.3.1. The pair $(V^A_{(u_*,\nu_*)}(\cdot),V^B_{(u_*,\nu_*)}(\cdot))$ represents a *fixed point* of the recursion (2.11)-(2.12) with boundary conditions (2.13)-(2.16).

Given the fixed point $(V^A_{(u_*,\nu_*)}(\cdot),V^B_{(u_*,\nu_*)}(\cdot))$, the solution of (2.11) and (2.12) allows us to characterize the optimal pair of investment rules (u_*,ν_*) for every state (n,m,δ) as follows[11]

$$u_*(n,m,\delta)=\begin{cases} 1 & \text{if} \quad \pi^A E_\delta^Q[V^A_{(u_*,\nu_*)}(n+1,m,\tilde{\delta})-V^A_{(u_*,\nu_*)}(n,m,\tilde{\delta})]>a+b\delta \\ 0 & \text{if} \quad \pi^A E_\delta^Q[V^A_{(u_*,\nu_*)}(n+1,m,\tilde{\delta})-V^A_{(u_*,\nu_*)}(n,m,\tilde{\delta})]<a+b\delta \end{cases}, \quad (2.17)$$

[11]Since the process δ is a continuous-state process, there is no loss in generality by considering only strict inequalities in (2.17) and (2.18).

and

$$\nu_*(n,m,\delta) = \begin{cases} 1 & \text{if} \quad \pi^B E_\delta^Q[V^B_{(u_*,\nu_*)}(n,m+1,\tilde{\delta}) - V^B_{(u_*,\nu_*)}(n,m,\tilde{\delta})] > a + b\delta \\ 0 & \text{if} \quad \pi^B E_\delta^Q[V^B_{(u_*,\nu_*)}(n,m+1,\tilde{\delta}) - V^B_{(u_*,\nu_*)}(n,m,\tilde{\delta})] < a + b\delta \end{cases} . \tag{2.18}$$

Note that since the state space is continuous (i.e., $(n, m, \delta) \in \mathbb{N} \times \mathbb{N} \times \mathbb{R}_+$), only pure strategies (u_*, ν_*) are played in equilibrium (*P-almost surely*). Moreover, since only one of the inequalities in (2.17) (and (2.18)) can be true in every subgame, the equilibrium is *unique* for every (n, m, δ).

From conditions (2.17) and (2.18), we note that the investment decision rule at every stage of the game is a simple NPV rule in which an investment step is undertaken if the expected marginal gain from a successful completion of the next stage is larger than the marginal cost, $a+b\delta$. Finally, because the optimal investment decision of a player is independent of the decision of the opponent, each player has a *dominant* strategy in every sub-game. It is interesting to emphasize, however, that this does not imply *independence* between the two players. In fact, the investment decision of a firm is dependent (through the value function) on the *position* occupied by the competitor, therefore, the strategy of one firm is *indirectly* affected by the strategy of the opponent, thus making the interaction between players' strategies a non trivial one.

In the remainder of this section we discuss some properties of the value function and investment strategies. In particular we provide a sufficient condition for the existence of threshold strategies (single-crossing property) and a characterization of the equilibrium value function.

2.3.1 Analysis of the Model

We begin by defining the concepts of *symmetric race* and of *neck-and-neck subgame.* Both will be used extensively in my analysis.

Definition 2.3.3 *A* symmetric race *is defined as a race in which the two competing firms have identical technologies, i.e. the same cost structure* (a, b) *and the same research skills* ($\pi^A = \pi^B \equiv \pi$).

Definition 2.3.4 *A* 'neck-and-neck' *subgame is a subgame within a race in which both firms have completed the same number of stages, i.e.* $n = m$.

An immediate property that holds for symmetric races is the following

Proposition 2.3.5 *In symmetric races (i.e.* $\pi^A = \pi^B = \pi$*), the only equilibrium strategies in neck-to neck subgames are* $[I, I]$ *and* $[W, W]$.

The intuition behind the above result is straightforward. In a symmetric, neck-and-neck game the two firms are indistinguishable ($n = m$ and $V^A(\cdot) = V^B(\cdot)$). Given the structure of the optimal investment strategy in (2.17)-(2.18) it is immediate to see that they decide to play *symmetric* strategies.[12]

For the remainder of the paper, we will consider symmetric races only.[13] In such cases, we can invoke symmetry and write

$$V^B(n, m, \delta) = V^A(m, n, \delta) \equiv V(m, n, \delta). \tag{2.19}$$

Also, for notational convenience, we will omit the subscript (u_*, ν_*) when referring to a value function. Unless otherwise specified, the value functions are to be intended as derived under the optimal pair (u_*, ν_*) of Markov strategies.

The first property of value functions is a useful monotonicity result in the number of stages completed by the competing firms.

Proposition 2.3.6 *The value of the R&D venture $V(n, m, \delta)$ is non-decreasing in n, non-increasing in m, and tends to zero as δ goes to zero, i.e.,*

(i) $V(n+1, m, \delta) \geq V(n, m, \delta)$, *for all* m, δ;

(ii) $V(n, m, \delta) \leq V(n, m+1, \delta)$,*for all* n, δ.

(iii) $V(n, m, \delta) \to 0$, *as* $\delta \to 0$ *for all* n, m.

As we would expect, the successful overcoming of a stage has a positive effect on the value of the successful firm and a negative effect on the value of the rival. In Section 2.5 I quantify these effects.

We now characterize the investment strategies emerging from the solution of the patent race model. A useful property that we require in the investment strategies is described in the following

[12] It is well-known that, if the set of actions of every player is convex, then a *symmetric* game has a *symmetric* Nash equilibrium (see, for example, Osborne and Rubinstein (1994), p. 20). However, this is not generally true if the action space is finite (as it is in our case). The "Hawk-Dove" game provides a counterexample. Our result is driven by ruling out the event of contemporaneous success when both firms invest.

[13] The solution technique we propose can be easily extended to non-symmetric races.

Definition 2.3.7 *A firm is said to follow a* threshold investment strategy, *if every decision rule in the strategy can be described uniquely in any given state* (n, m) *by a level of cash flow* $\delta_*(n, m)$ *such that mothballing occurs for* $\delta < \delta_*(n, m)$ *and investment occurs for* $\delta \geq \delta_*(n, m)$.[14]

Denote by $\Xi_{(n,m)}(\delta)$ the marginal gain/loss incurred by firm A when it faces an opponent at stage m and overcomes the $n+1$-th hurdle, i.e.

$$\Xi_{(n,m)}(\delta) = V(n+1, m, \delta) - V(n, m, \delta), \text{ for all } n, m. \tag{2.20}$$

The following proposition provides a sufficient condition for the existence of threshold investment decision rules.

Proposition 2.3.8 (Single-Crossing Property) *If* $\frac{\partial}{\partial \delta} E^Q[\Xi_{(n,m)}(\tilde{\delta})|\delta] \geq b/\pi$ *for all* n, m, *then there exists a level of cash flow* $0 < \delta_*(n, m) \leq \infty$ *such that firm* A *mothballs for* $\delta < \delta_*(n, m)$ *and invests for* $\delta \geq \delta_*(n, m)$.

If we restrict ourselves to the class of threshold investment strategies, then it is possible to show (see Proposition 2.3.9) that if a firm is leading the race (i.e. it has completed more phases than its rival), it tends to exercise its next development option at a lower threshold than the rival. We interpret this result by saying that *the leader in a race is more eager to invest than the follower.*[15]

Proposition 2.3.9 *If the conditions in Proposition 2.3.8 are satisfied and* $\Xi_{(n,m)}(\delta) \geq \Xi_{(m,n)}(\delta)$ *then, when* $n \geq m$, $\delta_*(n, m) \leq \delta_*(m, n)$.

For an intuition behind this result we can rely on an analogy from option pricing theory. Note that $V(n, m, \delta)$ is the value of a compound option with exercise price equal to the expected investment cost to completion. If $n \geq m$ than the compound option $V(n, m, \delta)$ is less leveraged than $V(m, n, \delta)$. The condition $\Xi_{(n,m)}(\delta) \geq \Xi_{(m,n)}(\delta)$ requires that the marginal gain from overcoming hurdle $n+1$ (when the opponent is at stage m) is higher than the marginal gain of overcoming hurdle $m+1$ (when the opponent is at stage n). This may happen when the option $V(n, m, \delta)$ is closer to be "at the money" than the option $V(m, n, \delta)$.

[14] *We used the convention that investment occurs for level of cash flows greater* or equal *to* δ_*. *Given that* δ *is a continuous-state process this is without loss of generality.*

[15] Grossman and Shapiro (1987) find an analogous property. In their paper they show that the leader in a race "works harder" than the follower.

We conclude this section by showing a striking feature of the value function emerging from the solution of the system (2.11)-(2.12). Under some circumstances, in fact, the value function may not be continuous in δ. The following proposition shows when this may happen and quantifies the size of the discontinuity.

Proposition 2.3.10 *If the conditions in Proposition 2.3.8 are satisfied, then*

(i) The value function $V(n, m, \delta)$ solving (2.11)-(2.12) can be characterized as follows

$$V(n, m, \delta) = V_1(n, m, \delta)\chi_{\{\delta<\delta_*(m,n)\}} + V_2(n, m, \delta)\chi_{\{\delta\geq\delta_*(m,n)\}}$$

where V_1 and V_2 are continuous, non-decreasing functions of δ and χ is the indicator function. Moreover, $\lim_{\delta\to\delta_*^-} V_1(n, m, \delta) \geq \lim_{\delta\to\delta_*^+} V_2(n, m, \delta)$.

(ii) $\lim_{\delta\to\delta_*^-} V_1(n, m, \delta) - \lim_{\delta\to\delta_*^+} V_2(n, m, \delta) \leq \pi\frac{\beta\delta_*}{(1-\beta)^2}$. i.e. the upper bound of the discontinuity in $V(n, m, \delta)$ at δ_* is given by $\pi\frac{\beta\delta_*}{(1-\beta)^2}$.

The intuition behind the above result is simple. $V_1(\cdot)$ can be interpreted as the value function of a firm facing an opponent who always waits, while $V_2(\cdot)$ is the value for a firm facing an opponent who always invests. If the single crossing property (Proposition 2.3.8) holds, then the strategies played by the two firms are threshold strategies. This justifies $V(\cdot)$ being a mixture of $V_1(\cdot)$ and $V_2(\cdot)$. Moreover, by Proposition 2.3.6, facing a rival who always waits is at least as good as facing a rival who always invests (the rival in the former case never progresses). This explains the nature of the discontinuity of $V(\cdot)$ in δ_*. Recalling that the probability π is a shorthand notation for $\pi\Delta$, it is evident from (ii) that the discontinuity will disappear in continuous time, as $\Delta \to 0$.

2.4 The Cooperative Outcome

In this section, we construct a benchmark to evaluate whether, and to what extent, the risk emerging from preemption in technological races contributes to the determination of risk premia earned by the competing R&D ventures. As in Grossman and Shapiro (1987), this benchmark is represented by a *Research Joint Venture* (RJV), defined as an agent who has at its disposal two research units, identical to the ones competing in R&D. The problem faced by this agent is to optimally coordinate investments in those units in order to maximize the aggregate value. We emphasize that such a benchmark is a purely theoretical construct; it serves the purpose of defining an idealized player who does not face preemption in the development process. Acting in the absence of strategic incentives, the investment policies and, consequently, the values of the research units managed by the RJV represent a *non-strategic* solution to the investment problem. In order to investigate the impact of strategic interactions on value and risk, we will compare these variables with the corresponding quantities derived from the solution to the patent-race problem.[16]

Let $V^J(n,m,\delta)$ indicate the value function of the RJV when unit 1 has completed n stages and unit 2 has completed m stages. No synergies are enjoyed by the RJV as a result of a joint research effort. This is done in order to have a homogeneous benchmark for comparison and implies that the technology is not affected by the existence of a RJV. The action space, cost and technology of the RJV are summarized in Table II. Following familiar steps, the value $V^J(n,m,\delta)$

Table II. The RJV problem

Action	*Cost*	*Prob. of success*
mothball both R&D units	0	0
Invest in R&D unit #1	$a+b\delta$	π
Invest in R&D unit #2	$a+b\delta$	π
Invest in both R&D units	$2(a+b\delta)$	2π

of the coordinated R&D project satisfies the following Bellman equation

$$\begin{aligned}
V^J(n,m,\delta) &= max\Big\{E_\delta^Q[V^J(n,m,\tilde{\delta})], \\
&\quad \pi E_\delta^Q[V^J(n+1,m,\tilde{\delta})] + (1-\pi)E_\delta^Q[V^J(n,m,\tilde{\delta})] - a - b\delta, \\
&\quad \pi E_\delta^Q[V^J(n,m+1,\tilde{\delta})] + (1-\pi)E_\delta^Q[V^J(n,m,\tilde{\delta})] - a - b\delta, \\
&\quad \pi E_\delta^Q[V^J(n+1,m,\tilde{\delta})] + \pi E_\delta^Q[V^J(n,m+1,\tilde{\delta})] + \\
&\quad (1-2\pi)E_\delta^Q[V^J(n,m,\tilde{\delta})] - 2(a+b\delta)\Big\} \qquad (2.21)
\end{aligned}$$

with boundary conditions

$$V^J(N,m,\delta) = \frac{\delta}{1-\beta},\ m<N; \quad V^J(n,N,\delta) = \frac{\delta}{1-\beta},\ n<N. \qquad (2.22)$$

[16]An alternative way of looking at my benchmark is by thinking of it as a case of *extreme collusion* in the R&D process. Therefore, by comparing values and risk premia from the R&D game to the corresponding quantities in the RJV problem, we are implicitly comparing across investment processes characterized by different degrees of competitive intensity.

The four arguments inside the *max* operator of equation (2.21) represent respectively the decision to moth ball both units ($[W, W]^J$), invest in R&D unit #1 ($[I, W]^J$), invest in unit #2 ($[W, I]^J$) and invest in both units ($[I, I]^J$). The solution to the above infinite horizon Bellman equation is relatively straightforward since no strategic issues are involved. In the following proposition we show that (2.21) uniquely defines a *fixed point* $V^J(n, m, \delta)$.

Proposition 2.4.1 *The Bellman equation (2.21) admits a unique fixed point $V^J(n, m, \delta)$, for all $n = 0, \dots, N$, $m = 0, \dots, N$.*

We conclude this section by reporting some of the properties of the benchmark RJV solution.

Proposition 2.4.2 *The function $V^J(n, m, \delta)$ solving the coordinate R&D problem,*[17]

(i) is monotone and convex in δ;

(ii) is increasing in n and m;

(iii) approaches zero as δ approaches zero.

As one may expect, the RJV, not facing the fear of preemption in the investment process, is *more patient* in undertaking development. In our context, 'more patient' means that the investment is triggered at a higher level of cash flows, compared to the patent race case. The next proposition formalizes this concept and provides a sufficient conditions for this to happen. Let us first define the quantity

$$\Xi^J_{(n,m)}(\delta) \equiv V^J(n+1, m, \delta) - V^J(n, m, \delta). \tag{2.23}$$

$\Xi^J_{(n,m)}(\delta)$ represents the marginal gain/loss faced by the RJV when it is successful in overcoming the $n+1$-th hurdles in unit #1. This is the non-strategic equivalent of equation (2.20).

Proposition 2.4.3 *If $\Xi^J_{(n,m)}(\delta) \leq \Xi_{(n,m)}(\delta)$ for all δ, then there exist two thresholds $\delta_*(n, m) \leq \delta^J_*(n, m)$ such that firm A invests for $\delta \geq \delta_*$ and the RJV invests in unit #1 for $\delta \geq \delta^J_*$.*

The condition stated in the above proposition is rather intuitive. If the marginal gain that the RJV obtains from overcoming a hurdle is less than the gain that a competitor obtains from winning a stage at the expense of the opponent, then investment in a particular R&D unit occurs "later" if it

[17] *Note that, by construction, $V^J(n, m) = V^J(m, n)$ for all n, m.*

is undertaken by a RJV rather than if it was undertaken by an independently competing R&D unit. This is because the RJV is indifferent (by construction) between which research unit is successful, while it is obviously not the case for the single competing venture.

Proposition 2.4.3 states a condition under which my model predicts the (empirically) well-known result of excess of R&D spending in industries with competitive pressures. A natural consequence of this proposition is that *the expected time to completion for the winner of a patent race is lower than the expected time to completion for the RJV.*

2.5 Numerical Results

In this section, we solve numerically for a stationary equilibrium in the patent race model and compare the resulting value functions, investment strategies and risk premia to the corresponding quantities emerging from the collusive outcome. The numerical solutions allow me to study the impact of technological competition on values and returns. In particular, we analyze whether competition in R&D is responsible for value-dissipation and how the risk of preemption in a technological race may contribute to the overall riskiness of the competing R&D firms. Details of the numerical procedure used to solve the models are described in Appendix B.

We calibrate the model to quarterly data, i.e. one "period" corresponds to three month.[18] The parameters used in our numerical analysis are described in Table III. We consider an R&D project that requires $N = 20$ phases to be completed. Since we use quarterly data this means that, if a firm is always successful when it makes an investment, and if we ignore the possibility of a rival winning some stages, it will take the firm 20 quarters (i.e. 5 years) to finish and commercialize the product of R&D. We assume a fixed cost $a = .5$ per period and a variable cost $b = .05$ per period. Therefore, for values of δ around 10 the fixed and variable costs contribute equally to the total cost. The probability of technical success is assumed equal to 25% for both players. This refers to the chance of being successful in a quarter if investment is undertaken.[19] The *potential* cash flow process (δ) grows at an average of 3% per year ($\mu = 0.0075$) with an annual standard deviation of about 60% ($\sigma_\delta = 0.3$). We select the standard deviation, σ_z, of the pricing kernel and

[18]As discussed in Section 2.2, in such a period it is relatively unlikely that two research labs are equally successful in the research process.

[19]Note that we are solving a *symmetric* version of the game, according to Definition 2.3.3.

the correlation, ρ, between the pricing kernel's innovations and those of the cash flow process in order to obtain a price of risk $\lambda = 12\%$ (p.a.). This is also the risk premium demanded by the completed project as we show in Proposition 2.5.3. Since the (real) risk free rate is assumed to be 7%, this implies an expected return of 19% for the completed project.[20]

2.5.1 The Effect of Preemption on Value

To assess the impact of preemption on value we first analyze some of the properties of the value functions $V(n, m, \delta)$ emerging from my numerical solution of the patent race model. We then compare those values across industry structures to quantify the dissipation effect of technological competition.

Table IV reports, for different values of cash flows, the value functions in three different life-stages of the development process: early stage (Panel A), intermediate stage (Panel B) and late stage (Panel C).[21] The following result describes an interesting property of value functions.

Result 1 *Successes and failures have* asymmetric *effects on the value of competing R&D ventures.*

Consider, for example, the case of a neck-and-neck race at stage (10,10) (i.e. in the intermediate phase of development). If the current level of cash flow is $\delta(t) = 20$ and firm A wins the first stage at the expenses of B, its value jumps from $V(10, 10, 20) = 42.097$ to $V(11, 10, 20) = 65.281$, an increment of about 55%. On the other side, losing that stage has a relatively mild effect, reducing the value to $V(10, 11, 20) = 32.473$, an overall loss of about 23%. The asymmetry is more pronounced for lower value of cash flows where the convexity of value functions is higher. The explanation for this result relies on two crucial points: (1) the infinite horizon nature of the game and (2) the winner-take-all provision. Because (i) the game is an infinite horizon game, (ii) waiting is costless, and (iii) there is no learning about research skills of the opponent,[22] it is never optimal for a player to abandon the race. The value of the loser is therefore supported by a substantial "option to catch up" that counter-balances the loss in value due to failures. This property is of course shared *ex-ante* by the winner who, in addition, reduces its distance from the completion of

[20]Robustness of all the results reported in the remainder of the paper have been extensively checked against different parametrizations.

[21]We numerically verify that the single crossing condition in Proposition 2.3.8 is satisfied in all simulations.

[22]The cost structure and the probability of success are common knowledge.

the project. From Proposition 2.3.6, this has a positive effect on value. A second interesting result is the following

Result 2 *The marginal effect of a success is not constant among the different subgames.*

In particular, given the position m of the opponent and a level of cash flows, the function $n \to V(n; m, \delta)$ is convex in n. This property is responsible for the "eagerness of the leader"-property proved in Proposition 2.3.9 and is illustrated in Figure C.2, showing a family of value functions for firm A when B has completed $m = 8$ phases. we also noted that, looking at symmetric games and using the definition of marginal gain/losses in value (equation (2.20)), if $n \geq m$, $\Xi_{(n,n)}(\delta) \geq \Xi_{(m,m)}(\delta)$ for all δ. Therefore, by the optimal investment strategy (2.17) and (2.18), $\delta_*(n, n) \leq \delta_*(m, m)$. In other words, the exercise of development options, in neck-and-neck scenarios, occurs for lower levels of cash flow as the race progresses towards completion. Moreover, as the opponent draws even, both become more eager in the exercise of their investment options, i.e. $\delta_*(n, n) \leq \delta_*(n, m)$, $n \geq m$. This is also illustrated in Table IV where we indicate with '*' the decision of Firm A to moth-ball. *Mutatis mutandis*, these properties confirms some of the findings of Harris and Vickers (1985), Fudenberg et al (1983) and Grossman and Shapiro (1987) on the the intensity of rivalry within a race. In particular, Grossman and Shapiro (1987) claim that "when one firm is ahead and the other behind, the leader's speed exceeds that of the follower [and] the (former) follower increases its effort if it manages to draw even." (p. 377) Another consequence of this property on investment triggers is that the follower in a race is facing a higher leverage investment option, i.e. an investment opportunity with a higher expected time to completion. This is reflected in the effect on values documented above and on risk premia (see below).

So far, we have looked at values as a function of underlying cash flow and *absolute* position of firms in the technology race (i.e. number of phases completed). It would be interesting to see how value is affected by the *relative distance* between firms. How does the value of a firm who is ahead $n - m$ stages of its rival change as the distance to completion varies? To address such a question, we need to examine the problem from a different perspective. We interpret the relative distance to the rival as a measure of *intensity* of rivalry. The smaller (in absolute value) is $|n - m|$, the higher is the intensity of rivalry. According to this definition, neck-and-neck subgames, as defined in Definition 2.3.4, represent the highest degree of rivalry between firms engaged in a patent race.

Table V displays the value function for different degrees of rivalry and for different stages of the development process. Panel A ($n = 5$) refers to early development, Panel B ($n = 10$) to intermediate development and Panel C ($n = 15$) to late development. As it is evident from this table, for any given intensity of rivalry $(n - m)$, the value of the R&D venture is increasing as the venture approaches completion (i.e. as we move from Panel A to Panel C). Moreover, given an absolute distance to completion for firm A (i.e., fix n), the value decreases as the relative distance decreases (i.e., as m increases). These results are a trivial consequence of the monotonicity property shown in Proposition 2.3.6. More interesting is the effect on *marginal* gains/losses in value as a function of relative and absolute distance which we summarize in the following

Result 3 *Given a relative distance $(n - m)$ between firms, the value function is decreasing and concave in the absolute distance to completion.*

We can see this property by looking at the percentage marginal gain in value from successfully completing a stage for a given intensity of rivalry $(n - m)$. From Table V we notice that this is decreasing in the distance to completion. Consider, for example, the neck-and-neck subgames in Table V. When $\delta = 20$ a firm in intermediate development ($n = 10$) is worth 7 times more than a firm in early development (42.096/5.990) while a firm in late development ($n = 15$) is only approximately three times as valuable as a firm in intermediate development (123.914/42.096).

A comparative static exercise on the volatility of cash flow unveiled the following

Result 4 *The % marginal effect of a success/failure is increasing in the volatility of cash flows.*

We solved the model for $\sigma_\delta = .15$ and $\sigma_\delta = .10$ representing respectively the cases of medium volatility of cash flows (30% p.a.) and low volatility (20% p.a.). Table VI reports the value function in different stages of developments for different levels of volatility of cash flows. Note that the effect of a low volatility, *ceteris paribus* is to increase the final value of the completed project through a reduction of the risk-adjusted discounted factor β (See Proposition 2.3.1). Note also that the marginal effect of overcoming a hurdle is higher when uncertainty is higher. Consider for example the % gain in value of successfully completing stage 10 by firm A when $\delta = 20$. The marginal gain is 23.32% when $\sigma_\delta = .10$, 26.28% when $\sigma_\delta = .15$ and 55.07% when $\sigma_\delta = .30$. Similarly the % marginal loss of losing a stage to the opponent is higher for higher level of volatility (when $\delta = 20$, the marginal loss from losing stage 10 is 18.4% if $\sigma_\delta = .10$, 18.5% if $\sigma_\delta = .15$ and 33.9% if $\sigma_\delta = .30$). Moreover, by comparing the exercise strategies of firm A across different levels of volatility (See Table VI) I note that higher uncertainty in cash flow is also associated with higher exercise triggers.

In Proposition 2.3.10 we showed that the value function may be discontinuous in δ.[23] An example is given in Figure C.3. Panel A shows the value function in subgame $(8, 6)$, while panel B reports the investment strategies played by both firms. As we can see, the value function is discontinuous at $\delta_* = 15.50$, when, for the rival firm, it is optimal to invest in the attempt to catch up.

We now address the problem of identifying whether R&D competition has a value-dissipating effect by comparing the value of the cooperative outcome (RJV) with the total industry value emerging from a patent race. It can be useful to interpret the RJV value as the *industry value* of the R&D project, where the industry is represented by the RJV itself. Similarly, in the case of a competitive industry, we can define the industry value of the R&D project as the joint value of the competing firms. The comparison of these two quantities sheds light on the influence of R&D competition on market values of R&D projects. To perform this analysis, we introduce the following measure of *erosion in value.*

Definition 2.5.1 *Given the solution $V(n,m,\delta)$ of the patent race problem (2.11)-(2.12) and the solution $V^J(n,m,\delta)$ of the RJV problem (2.21), we define the percentage* erosion in value *as the quantity*

$$\Phi(n,m,\delta) = \frac{V^J(n,m,\delta) - [V(n,m,\delta) + V(m,n,\delta)]}{V^J(n,m,\delta)}.$$

$\Phi(n,m,\delta)$ is a measure of the percentage loss/gain in market value due to the presence of technological competition. If $\Phi(\cdot) > 0$ then competition is *value-dissipating.*

Table VII reports some of the percentage erosion in early stages of development (Panel A), intermediate stages (Panel B) and late stages (Panel C). All the entries in the table are non-negative, which supports the following

Result 5 *R&D competition is value-dissipating.*

[23] Recall that such a discontinuity is due to the discrete-time solution of the model and would disappear in continuous time. Since δ is monitored in discrete-time intervals, there is no possibility of constructing an arbitrage strategy that may exploit such a discontinuity.

We also observe that the percentage erosion is generally a decreasing function of the level of cash flows δ.[24] Moreover, notice that erosion tends to he higher in neck-and-neck subgames than in asymmetric subgames and, generally decreases with the number of phases completed.

What is the cause of value dissipation? Value (profits) are a reflection of optimal exercise policies. Since the investment strategies implemented by the RJV are exempt of strategic issues and the technology structure is unchanged, the difference between the industry value of RJV and the total industry value of competing R&D ventures must originate uniquely by the effect of preemption in the R&D race. To understand the causes of value-dissipation we need to look at optimal investment strategies in the patent race and in the RJV problem. In Proposition 2.4.3 we provided a sufficient condition for the RJV to be more patient (i.e. to exercise its development options later) than a competing venture. I numerically checked those conditions and found them satisfied in all subgames. Compared to the case in which firms merge, we observe that investment thresholds δ_* are lower in the patent race case. I interpret this finding as evidence of a higher intensity of R&D investments in the patent race.

Result 6 *R&D competition induces a higher intensity of R&D investments.*

We illustrate this result with an example. Table VIII reports the value functions and exercise strategies in the race and in the RJV problem in subgames (8,0) and (8,8). The first subgame is a "quasi-monopoly" case, since firm A is facing a rival who has not developed yet. Note, indeed, that the value and strategy for firm A are identical to the one of a RJV. The second subgame is a "neck-and-neck" situation. Notice how, in this situation, the discrepancy in values is more dramatic, as summarized by the percentage erosion. The comparison between the investment strategies shows that the ventures in the race rush to develop while the RJV, not facing the risk of preemption, is more patient in the exercise of its options. This is responsible for the erosion in value. Since this fear of preemption is bigger when competitors are neck-and-neck, it is not surprising to observe, in these scenarios, higher values of erosion and an earlier exercise of development options.

The insight from the above analysis is that preemption induces rivals to exercise their development options earlier than it would be "optimal" in a no-preemption situation and this

[24]This depends on the fact that $V(n,m,\delta) + V(m,n) \to V^J(n,m)$ as $\delta \to \infty$. The proof of this relies on the fact that $V(n,m) + V(m,n) = V^J(n,m)$ when the fix-cost component is null ($a = 0$). In this case, in fact the value function is homogeneous in δ. We characterize this case in the next subsection.

is responsible for the erosion in values. The value-dissipating effect of competition is, however, counterbalanced by the fact that the expected time to completion is shorter (as a consequence of Proposition 2.4.3) for the winner of a race than for the RJV. This brings us naturally to the next question. How "bad" is preemption from the point of view of an investor who is about to invest in the competing firms? To answer this question we need to investigate how markets perceive risk in competitive and non-competitive markets.

2.5.2 The Effect of Preemption on Risk Premia

Risk premia summarize the views of the markets on the riskiness of the R&D project and indicate the proper discount rate to be used in evaluating future uncertain cash flows. Let us start by defining the required expected returns on an ownership claim to a competing R&D venture.

Let $R^i(n,m,\delta)$ denote the return on a ownership claim of firm $i = A, B$ which is currently engaged in a race at state (n,m,δ). Expected (gross) return on assets of firm $i = A, B$ in state (n,m,δ) are defined as

$$\text{Expected return}^i_{(n,m,\delta)} = \frac{E^{\Pi,P}_{(n,m,\delta)}[\text{Next-period Cum-dividend value}]}{\text{Current-period Ex-dividend value}}, \quad i = A, B,$$

where Π is the probability measure governing technical uncertainty (the transition probability matrix in Table I), P is the probability law of δ and $E^{\Pi,P}_{(n,m,\delta)}(\cdot)$ is the expected value conditional on the state (n,m,δ). Note that the only cash flow generated by the firms engaged in the patent race is a negative cash flow represented by the investment cost $I(\delta) = a + b\delta$.[25] The following definition formalizes the concept of expected return and risk premium for an R&D venture engaged in a patent race.

Definition 2.5.2 *Given the transition probability matrix* Π *(technical uncertainty) and the probability measure* P*, expected gross returns for firm* $i = A, B$ *are defined as follows*

$$E_{(n,m,\delta)}[1 + \tilde{R}^i] = \frac{E^{\Pi,P}_{(n,m,\delta)}[V^i(\tilde{n},\tilde{m},\tilde{\delta})]}{V^i(n,m,\delta) + \chi(n,m,\delta)(a + b\delta)}, \tag{2.24}$$

where $\tilde{n}$ *and* $\tilde{m}$ *are random variables evolving according to the probability law* Π*,* $\tilde{\delta}$ *evolves according*

[25] This is typical of R&D ventures.

to (2.1) and

$$\chi(n,m,\delta) = \begin{cases} 0 & \text{if} \quad \delta < \delta_*^i(n,m) \\ 1 & \text{if} \quad \delta > \delta_*^i(n,m) \end{cases}, i = A, B.$$[26]

The risk premium, $RP(n,m,\delta)$ is the excess expected return (in subgame (n,m)) over the continuously compounded risk-free rate:

$$RP(n,m,\delta) \equiv \ln\left(E_{(n,m,\delta)}[1+\tilde{R}]\right) - r. \tag{2.25}$$

Of course, in order to compute the expected return in (2.24) I need to keep track of the possible equilibria emerging during the game which will inevitably affect the expectation with respect to the probability of technical success (Π). For example, if $[I,W]_{(n,m,\delta)}$ is the equilibrium in state (n,m,δ) we would have

$$E^{\Pi,P}_{(n,m,\delta)}[V(\tilde{n},\tilde{m},\tilde{\delta})] = E^{P}_{(n,m,\delta)}[\pi V(n+1,m,\tilde{\delta}) + (1-\pi)V(n,m,\tilde{\delta})]$$

The other cases are dealt with similarly. In what follows, all the quantities we define are referred to player A, without loss of generality.

As for the case of value, we start our investigation on risk premia by first looking at the properties of returns obtained from the patent race model. The analysis of risk premia within a patent race will shed light on how the different sources of uncertainty (technical, systematic and "preemption-driven") contribute to the overall rate of return. We then design an appropriate way of comparing risk premia across markets characterized by different degrees of R&D competition.

Figures C.4 and C.5 plot the risk premium (% p.a.) demanded by a share of firm A against the number of phases completed (n), and the level of cash flows (δ), assuming that the opponent

[26] To clarify this definition, suppose that a firm in stage n, facing a competitor at stage m, needs $I(\delta) = a + b\delta$ dollars to make its investment in R&D. To raise such an amount, assume that the firm is forced to issue shares, since the riskiness of the project prevents any other form of financing. Given that the cum-dividend value at time t is $V(n,m,\delta)$, the *price* of each share (i.e. the ex-dividend value of the firm) is $V(n,m,\delta)+I(\delta)$, where the cum dividend part (V) can be interpreted as the "growth component" embedded in the share price. The number of shares issued to raise I dollars will be $I(\delta)/(V(n(t),m(t),\delta(t))+I(\delta))$ and the total cost for the financier will be, of course, $I(\delta)$. Now, the expected value at time t of the investment time $t+1$ will be

$$E^{\Pi,P}_{(n,m,\delta)}[V(\tilde{n},\tilde{m},\tilde{\delta})] \times \frac{I(\delta)}{(V(n,m,\delta)+I(\delta))}$$

and consequently, dividing by the initial investment I, we obtain (2.24) as the correct expression for expected returns.

has completed $m = 1$ and $m = 8$ stages respectively. From the figures we can roughly grasp some general properties of risk premia. The first obvious property is that

Result 7 *Risk Premia $RP(n, m, \delta)$ are not monotonic in δ.*

From Figure C.4 we observe that risk premia are generally higher for lower levels of cash flows, which usually corresponds to regions where it is optimal to mothball the project. Moreover, the risk premia seem to drop dramatically as the economic conditions (i.e. δ) improve. Recall that this corresponds to an incentive in the exercise of the investment option open to the players. These results are very similar to the one noticed by Berk, Green and Naik (1999). This is not surprising since, when firm B is at stage $m = 1$, we are basically looking at risk premia for a player (firm A) facing an opponent who is very unlikely to be a threat. Things change dramatically if we analyze risk premia of a venture facing a competitor who is reasonably far ahead in the development process. Figure C.5 shows the risk premia dynamics of firm A when the opponent has completed $m = 8$ stages. The most striking issue is the fact that risk premia can be increasing in δ, as it can be seen for relatively low levels of n. This increase in risk premia is due to the fact that the opponent is investing in this region, while firm A (to which the value of risk premia are referred to) is mothballing. Therefore, *when the opponent pulls ahead, the risk premium for the ventures who mothballs increases.* The risk premium drops as soon as economic conditions improve to the point that it becomes optimal for the lagger to follow up. This is in contrast with what Berk, Green and Naik (1999) observe in their analysis of risk premia. The presence of a rival changes the behavior of risk premia demanded by a firm in the mothballing region, since it changes the characteristic of the mothballing option held by the competing firms. In other words, the risk premia dynamics is not only affected by the firm's investment decision, but also by its rival's, even though this is a purely idiosyncratic factor.

Under the conditions specified in Proposition 2.3.8, we know that the investment strategies of firms are characterized by cutoff rules $\delta_*(n, m)$ on cash flows. This property has a crucial effect on risk premia. The decision to invest is done by taking into account three dimensions: (i) level of future cash flows, (iii) research skills (technical uncertainty) and (ii) position in the game. The dependence of the investment decision on δ imparts an *option value* to the R&D project itself. As far as this option is valuable, we would expect to observe a risk premium in excess of the risk premium demanded by a perpetual dividend stream $\delta(t)$. Moreover, investment thresholds are a function of the relative positions of the players and, indirectly, embed the fear of being preempted by the rival ("preemption risk"). In other words, decisions to continue R&D spending are made contingent not only on the resolution of systematic and unsystematic uncertainty, but also on the strategic positioning in the game. Given that technical uncertainty and strategic consideration contribute in the the determination of the threshold δ_*, we would expect to see these sources of uncertainties to be priced, even if purely idiosyncratic. Being the decision to invest based on the evolution of the underlying profitability of the project (systematic risk), every factor that alter δ_* will impact the value of the option to mothball and, consequently, we may have that purely idiosyncratic factors may contribute to the overall riskiness of a competing venture. The evolution of risk premia in Figures C.4 and C.5 suggests the following analogy with option pricing. If we interpret the expected cost to completion as the "leverage" of the investment option then, for the winner of a stage, overcoming an hurdle is equivalent to reducing the leverage of its option while for the firm who is preempted, losing a stage is equivalent of an increase in the leverage of its investment option.

We now compare the risk premia earned by an R&D venture in different subgames. Table IX reports the risk premia in early stages, intermediate stages and late stages of the development process. From this table it is immediate to observe the following property of risk premia

Result 8 *Risk Premia $RP(n, m, \delta)$ are non monotonic in n and m.*

Even though, in general, risk premia $RP(n, m)$ tend to be non-increasing in n and non-decreasing in m—suggesting that losing a phase to the advantage of a rival increases the riskiness of the venture—we do not have a monotonicity property for risk premia as the one we observed for value functions in Proposition 2.3.6. Note, for example, that $RP(10, 10)$ can be higher than $RP(0, 0)$ for $\delta \in [1, 4]$. This is due to the fact that, within this interval of cash flow values, the venture in subgame $(10, 10)$ is very likely to be preempted by its rival, losing stage 11 to the advantage of its competitor. This makes this venture *relatively* riskier (for $\delta \in [1, 4]$) than a venture in stage $(0, 0)$ which is not facing such a risk of preemption in this region of cash flows.[27] In other words, although we may be tempted to say that the higher the number of completed phases the lower the risk premium, this is not entirely true. Risk premia are a *relative* measure of riskiness in a race, and are significantly affected by the position occupied by players in the game. The fear of

[27]Ventures in subgame $(0, 0)$ are moth-balling for the whole range of cash flow reported in the table.

preemption at a later stage can impact pricing more than at an early stage, even if, overall, it may seem reasonable to claim that it is "riskier" to invest in earlier stages ventures. Notice, however that both such ventures demand a risk premium in excess of λ which confirms the above conjecture that purely idiosyncratic factors can have a systematic effect. Therefore, although purely idiosyncratic, preemption risk, as well as technical risks, are compounded in equilibrium to deliver a higher risk premium with respect to the premium demanded by the completed project.

An interesting property of the stochastic process used to model cash flows is that, under the assumed pricing kernel (2.3), a security that provides a dividend stream δ (i.e. the completed project) yields a per-dollar constant risk premium. I formalize this in the following proposition.

Proposition 2.5.3 *The per-dollar risk premium of the completed project is equal to* $\exp(\lambda)$.

The above proposition provides a useful benchmark to analyze whether and to what extent the "distance" from completion, technical uncertainty *and* the presence of a rival can affect risk premia. In Appendix A (Proposition A.0.1) I show that when the investment cost does not have a fixed component ($a = 0$) the value function is homogeneous in δ, i.e. there exists a function $h(n, m)$ such that

$$V(n, m, \delta) = h(n, m)\delta \text{ for all } n, m.$$

Note also that in the absence of fixed costs, the investment strategies are invariant to the level of cash flows. In other words, when $a = 0$, the investment strategies are not of the threshold-type, or, more accurately, the only thresholds are either $\delta_*(n, m) = 0$ (a firm always invests) or $\delta_*(n, m) = \infty$ (a firm abandons the project). This happens because, without fixed costs, there is no *option of waiting* associated with value. Technical uncertainty is resolved only through investment. If both investment cost and value are proportional to the underlying cash flow, once the project is suspended it will never be restarted, since both costs and benefits move proportionally. Hence, in the absence of fixed costs, the risk premium demanded by an R&D venture should not be different from the risk premium of the underlying cash flow. In other words, in the absence of fixed costs, there is no price for unsystematic uncertainty since the investment strategy is not dependent on the underlying potential cash flow δ. We confirm this conjecture in the following

Proposition 2.5.4 *In the absence of fixed costs, the per dollar continuously compounded risk premium demanded by a competing venture is constant and equal to* λ.

A direct consequence of the above proposition is that risk premia converge to λ as $\delta \to \infty$, as can be seen from Table IX. This occurs since for high δ the fixed component of the investment cost becomes negligible, the value function becomes homogeneous in the cash flows and the investment strategies are independent of δ.

We now look at risk premia of a venture as a function of absolute distance to completion of the project and and relative distance from the rival. The first measure captures the *pure progress* effect (i.e. distance to the finishing line), while the second is a measure of the *intensity of rivalry* in the race. Table X is the analogous of Table V for risk-premia. The most striking property of risk premia that emerges from Table X is the non-monotonicity in the relative distance between rival firms.

Result 9 *Risk Premia $RP(n, m, \delta)$ tend to be lower when the intensity of rivalry is high (i.e., $n - m$ is low).*

In fact, keeping the distance to completion fixed (pure progress effect), notice that, for every level of cash flows, the risk premia are lower the closer the two firms are. In Panel A and B (early and intermediate development) the lowest level of risk premia for firm A occurs when $n - m = 1$, i.e. when B is only one stage behind, while in late stages of development, risk premia are at the lowest level when A is leading by 5 units. *The intensity of rivalry is beneficial in keeping risk premia low.* In equilibrium, markets seems to demand higher risk premia for ventures that assume quasi-monopoly position in early stages of the technology race. This is because the presence of a competitor, although engaging the ventures in dissipative R&D expenditures, is responsible for a faster resolution of technical uncertainty and a quicker completion of the race. Risk premia seem to be greatly affected by this factor.

We try to confirm the above intuition by comparing risk premia emerging from the patent race with those emerging from the RJV benchmark described in Section 2.4. This implies a comparison of returns across two different markets characterized by different intensity of R&D rivalry. In one market, research activity is conducted cooperatively within a RJV (benchmark) and in the other, firms compete for a patent. The cash flows and the technological structure of these markets are otherwise identical. We assume that investors in both markets are well-diversified, holding in both cases the "market portfolio". Our purpose is to see whether these two markets differ substantially in the equilibrium risk premia they demand. For the case of competing firm the "market portfolio" is a value-weighted portfolio of the two competing firms as the following definition formalizes.

Definition 2.5.5 *Let $R^{pf}(n,m,\delta)$ be the return on the value-weighted portfolio of competing R&D firms when firm A has completed n stages, firm B has completed m stages and the underlying potential cash flow equals δ. The (gross) expected return on the value weighted portfolio is defined as follows*

$$E_{(n,m,\delta)}[1+\tilde{R}^{pf}] = \frac{E^{\Pi,P}_{(n,m,\delta)}[V(\tilde{n},\tilde{m},\tilde{\delta}) + V(\tilde{m},\tilde{n},\tilde{\delta})]}{V(n,m,\delta) + V(m,n,\delta) + \eta(\delta)(a+b\delta)}, \tag{2.26}$$

where where $\tilde{n}$ and $\tilde{m}$ are defined as in Definition 2.5.2 and

$$\eta(\delta) = \begin{cases} 0 & \text{if } \delta \in [W,W] \\ 1 & \text{if } \delta \in [I,W] \text{ or } [W,I] \\ 2 & \text{if } \delta \in [I,I] \end{cases}.$$

The risk premia on the portfolio is similarly defined as the excess return over the continuously compounded risk-free rate,

$$RP^{pf}(n,m,\delta) \equiv \ln(E[1+\tilde{R}^{pf}]) - r.$$

Similarly, the expected returns and risk premia earned by the RJV are defined as follows

Definition 2.5.6 *Let $R^J(n,m,\delta)$ be the return for a RJV in which unit #1 has completed n stages, unit #2 has completed m stages and the underlying potential cash flow equals δ. The (gross) expected return on the RJV is defined as follows*

$$E_{(n,m,\delta)}[1+\tilde{R}^J] = \frac{E^{\Pi,P}_{(n,m,\delta)}[V^J(\tilde{n},\tilde{m},\tilde{\delta})]}{V^J(n,m,\delta) + \xi(a+b\delta)}, \tag{2.27}$$

where $\tilde{n}$ and $\tilde{m}$ are random variables evolving according to the probability law Π, $\tilde{\delta}$ evolves according to (2.1) and

$$\xi = \begin{cases} 0 & \text{if } \delta \in [W,W]^J \\ 1 & \text{if } \delta \in [I,W]^J \text{ or } [W,I]^J \\ 2 & \text{if } \delta \in [I,I]^J \end{cases}.$$

The Risk Premium $RP^J(n,m,\delta)$ demanded in state (n,m,δ) is defined as the excess expected return over the continuously compounded risk-free rate:

$$RP^J(n,m,\delta) \equiv \ln(E^P[1+\tilde{R}^J]) - r.$$

Table XI reports the risk premia for the portfolio of competing firms and for the RJV in different stages of development. Panel A analyzes the case of a "quasi-monopoly", i.e., the case in which one of the firms (units) is far ahead of the rival in the technology race. Panel B consider the case of neck-and-neck firms (units). The first feature we notice from the table is the equivalence between competitive R&D effort and RJV in the quasi-monopoly case. I observed earlier (see Table VIII) that, in these situations, the value and investment strategies of the leader in the race and of the RJV are identical. It is not surprising therefore to find equivalence in the risk premia they demand in equilibrium. Panel B provides interesting insights on the effect of preemption on equilibrium risk-premia. In neck-and-neck situations, when the competitive pressure is high, we observe regions of cash flow for which the equilibrium risk premia is indeed lower than the risk premia demanded by a RJV. Considering that the latter represents a first-best result, this finding is of particular relevance. Portfolios of competing firms tend to demand a lower risk premium in early stages of development (subgame (5,5) in Table XI). For intermediate and later stages, the risk premia of the RJV are uniformly higher than those demanded by the portfolio of competing ventures. Note, however, that in subgame (5,5) not all the risk premia differentials are negative. For values of cash flow around $\delta = 8.0$ the portfolio of competing ventures indeed demands a *higher* risk premium than the RJV. This happens because $\delta = 8.0$ is the investment trigger, i.e., the the investment option is *at the money.* As we observed in subsection 2.5.2, this generates a rise in the value of risk premia demanded in equilibrium.[28]

The second interesting feature we can distill from Table XI concerns the magnitude of the differentials in risk premia.

Result 10 *Despite the magnitude of value-dissipation due to preemption, risk premia demanded by the RJV are only slightly lower than risk premia demanded by a value-weighted portfolio of competing R&D ventures.*

The highest positive differential between R^{pf} and Rp^{J} I observe in our numerical simulation is 3.9% (subgame (4,4)) and the highest negative differential is 2.43% (subgame (4,4)). In both of these subgames, we observe an erosion in value as high as 50% of the RJV value. This suggests that, from the point of view of an outside investor, having the opportunity to invest in competing R&D ventures is not necessarily a "bad" thing. The most likely interpretation of these results

[28]It may be useful to think of the elasticity of an at-the-money option.

relies on the interaction between *pure progress effect* and *rivalry effect* on risk premia. As we showed in subsection 2.5.1, the RJV is more patient than a venture in a race. Once we combine competing ventures in a portfolio, I eliminate, *from the point of view of the outside investor*, the rivalry effect and are left with a process (the combination of the two competing ventures) that leads to the finishing line quicker than the RJV. Remember that by Proposition 2.4.3 the expected time to completion for the leader in a patent race is lower than the expected time to completion for the RJV. Therefore, what we previously interpreted as "sub-optimal" investment decisions, responsible for value-dissipation, now becomes a desirable feature for an investor who sees the ventures in its portfolio more rapidly approaching completion. This may explain why the early stages of development are judged less risky if investments are undertaken in a purely competitive environment than centralized through a RJV. When the project is close to completion, there seems to be a slight preference for avoiding dissipative competition induced by the rivalry effect.

The main insight from this analysis is that, although there is no clear and uniform relationship between risk premia across different industry structures, *R&D competition in early stages contributes to lower the systematic risk of the portfolio of competing ventures, despite the significant value-dissipation.* This confirms the relationship we found earlier between risk premia and intensity of rivalry (see Result 9). A well-diversified investor looks with interest at "sub-optimal" investment strategies that have the effect of speeding up the process of innovation. This is indeed what happens in the patent race model, as we showed in Section 2.3.1. Our analysis suggests also that the *speed* at which innovations are marketed may actually underly these results. It would be very interesting to formally investigate the effect of expected time to completion on value and risk premia. We believe my model is rich enough to address this question.

2.6 Conclusions

In order to assess the pricing implication of technological competition, we develop a capital budgeting model for an R&D firm engaged in a patent race. We model the patent race as a stochastic, non-cooperative game with two sources of uncertainty (technical uncertainty and systematic uncertainty). This allows us to embed the capital budgeting problem in an asset pricing framework, more suitable to address the main question of the paper.

Our analysis shows that the value of an R&D firm in a race reacts asymmetrically to successes and failures and that risk premia are dramatically affected by losing a stage in the development process. Moreover, risk premia are generally higher in the early stages of development and lower when the firms are closer to each other in the race. The comparison of value and risk premia across different industry structures unveils an interesting property of returns: in early stages of development, risk premia from the collusive outcome are generally higher than those demanded by a portfolio obtained by combining the competing ventures. The opposite is true in later stages of development. We find this result particularly interesting especially if we consider the substantial value-dissipation generated by R&D competition.

Chapter 3

Portfolio Selection with Multiple Assets and Capital Gains Taxes

3.1 Introduction

Capital gains taxes represent a crucial concern in the determination of optimal investment and consumption decisions over time. Despite the increasing presence in financial markets of tax-exempt and tax-deferred investors such as pension funds and mutual funds, fully taxable investors still represent the majority of market participants. According to the US tax code, taxes are incurred on gains or losses earned on securities only when those gains or losses are realized through the sale of the assets. This endows the investor with a valuable tax-timing option. This tax-timing option will be traded off with optimal diversification in the the solution of the portfolio problem.

Costantinides (1983) shows that, if we allow the investor to sell short and fully use the proceeds of the short sale, the optimal investment policy can be separated from the optimal liquidation policy. The investor can obtain full use of the value of a position with an embedded capital gain without voluntarily realizing the gain. This argument holds true provided we can easily circumvent wash-sale rules imposed by the tax code.[1] In contrast, when short-sale and wash-sale constraints are in place, the "optionality" in the realization of capital gains taxes introduces a natural trade-off in the portfolio problem. Although capital gains taxes can be deferred by not selling, this, on the

[1]According to the US tax code it is not possible to deduct a loss from a sale of a security if the same security or "substantially identical" securities are purchased within a window of 60 days centered around the date of the sale.

other side, may generate significant departure from the optimal diversified portfolio obtained in the absence of taxes.

The purpose of this paper is to analyze the effect of capital gains taxes on optimal portfolio decisions of an investor who can trade in multiple risky securities. The effects of capital gains taxes on optimal diversification can be gauged in a meaningful way only in a multi-asset portfolio problem. Given the complexity that arises in solving for the optimal trading strategy in this context, we confine ourselves to the case of two risky assets. Our model builds on the analogy between taxes and transaction costs. On the surface, taxes can be thought of as a transaction cost upon selling. At a deeper level, however, it is evident that these transaction costs have a highly complex nature. Firstly, selling an asset at a loss can actually generate transaction revenue in the form of negative taxation. Secondly, and most importantly, the magnitude of the cost or benefit is itself dependent on the state variables of the problem (i.e., the portfolio holdings and the the bases of the assets). This makes the portfolio problem completely path dependent and extremely hard to solve. We make the problem tractable by modeling the cost basis as the weighted average of the purchase prices of the assets. Since, in a three asset (two risky, one riskless) economy, there is little role for substantially identical securities that can be used to disguise wash-sales, we closely model the tax code by ruling out the possibility of wash-sales. By doing this we can model the evolution of the basis without relying on exogenously specified optimality in the realization policy.

We analyze the impact of capital gains taxes on portfolio allocation for two classes of investors. The first class is composed of *sophisticated* investors who can trade individually in the separate securities. Alternatively, these investors can be though of as investing in tax-efficient mutual funds whose managers fully internalize the tax and diversification preferences of the investors. The second class is composed of *naïve* traders who rely on institutional investors (Mutual Funds and Index Funds) to allocate their savings. The solution of the portfolio problem for the sophisticated investor allows us to study the property of the optimal trading strategy in the presence of taxes and the certainty equivalent cost/benefit of capital gains taxes. The solution of the portfolio problems for the naïve investor is useful to asses the value of tax-timing flexibility and the tax effect of mutual fund turnover. We summarize hereafter the major results of our analysis

For the case of sophisticated investors, the optimal trading strategy is a function of the tax status of the traded assets. In particular we show that there are regions where investors choose not to trade in an asset which has an embedded gain. This property is analogous to what we would find in the solution of a fixed transaction cost problem. When wash-sales were allowed, the investor optimally sells the whole position in the risky asset with an embedded loss, realizes the loss and then rebalances the portfolio to the optimal portfolio that will emerge in absence of taxation. In doing so, optimal diversification is achieved and the cost basis is reset.

Next, we analyze how the holdings of one risky asset can affect the optimal trade in the other risky asset. Trading for tax motives is relatively more important when the correlation between assets is high, while trading for diversification motives dominates when the assets are negatively correlated. We also consider the case in which both assets are subject to taxation on the accrued gains and losses each period and compare it with the case in which the investor can optimally realize gains and losses separately. We construct a measure of the value of the tax-timing option in terms of certainty equivalent cost required to make an investor who pays taxes on accrual as well off as an investor paying taxes on realization. We show that such value is decreasing in the correlation between assets, in the volatility of the risky asset and in the degree of risk aversion of the investor. We provide an interpretation of these results in terms of option pricing theory where the volatility and risk aversion are part of the exercise price of the tax-deferral option.

To further investigate the role of the tax-timing option and diversification we look at the portfolio problem of naïve investors who can only trade in one of two types of funds. The first is an open-ended mutual fund which guarantees a pre-specified balance between risky assets and distributes turnover-related dividends to its shareholders. The second fund is an index fund which provides the same balance of risky assets as the mutual fund, but does not distribute dividends. Trading in a mutual fund is always dominated by trading in the two constituent asset, i.e., a tax-efficient fund dominates a high turnover mutual fund. The difference between the two scenarios is highest for those states in which the two assets have substantial embedded gains, i.e., when the tax deferral options are more valuable.

We also look at the tax consequences of mutual fund turnover by examining the certainty equivalent costs of trading in an index fund that is taxed like a single asset. When compared to the case of a mutual fund, the index fund allows less costly diversification, and is therefore preferred in cases where the asset holdings diverge from the target ratio. The mutual fund dominates in cases in which rebalancing motives require high trades of assets with embedded losses.

Finally, in order to more closely model the tax code, we impose a wash sale constraint which prevents the investor from selling and immediately repurchasing the asset for the purpose of resetting the tax basis. We observe that the trading strategy is dramatically altered in the case of an embedded loss due to the fact that non-convexity in the feasibility set generates discontinuities in the optimal investment strategy.

There is an extensive body of literature that deals with the portfolio allocation problem in the presence of capital gains taxes. Constantinides (1983, 1984) pioneered the study of optimal investment and liquidation policy in the presence of capital gains taxes. Dammon, Dunn and Spatt (1989) quantitatively assess the value of the tax-timing option by explicitly considering the different tax treatment of long and short-term capital gains. Dammon and Spatt (1996) investigate theoretically the optimal trading and pricing of securities subject to asymmetric taxation in the context of a no-arbitrage model. Contrary to the approach in the above papers, we analyze the trade-off between optimal diversification and tax costs of trading in the context of a dynamic intertemporal portfolio problem. Dybvig and Koo (1996) solve the portfolio problem with capital gains taxes and a single asset by explicitly keeping track of the history of the basis in the asset. They develop and implement numerical algorithms for the solution of such a problem. The complexity of their set-up allows for the solution of a maximum of four time periods.

The papers that are most closely related to ours are Dammon, Spatt and Zhang (2001) and Leland (2000). Dammon, Spatt and Zhang (2001) analyze the optimal dynamic consumption, investment and liquidation policies for the case of a single risky asset. Their focus is on analyzing the investment and consumption decision over the investor's lifetime in the presence of capital gains taxes and short-selling constraints. We extend their model to the case of multiple assets and, as discussed earlier, we do not allow for wash-sales in our model. Moreover, we do not focus on lifetime consumption behavior of a taxable investor, but concentrate instead on the trading policy of an investor who trades off diversification benefits and tax cost of trading. Moreover, in analyzing the investment problem with a mutual fund, index fund and capital gains taxes, we can assess and quantify the tax-timing and flexibility options as well as the tax effect of mutual fund turnover.

Leland (2000) examines the optimal implementation strategy of a mutual fund who wishes to maintain assets in exogenous fixed proportions. The objective function is specified as a "loss function" which keeps tracks of the deviations from the target portfolio. Leland (2000) explicitly solves for the no-trade region in the presence of capital gains taxes and transaction costs in the one-asset case and characterizes the no trading region with transaction costs only for the two-asset case. In solving the problem in the presence of capital gains taxes, Leland (2000) relies on simulation to determine the average cost bases as a function of the no-trade interval and uses this in the boundary condition of the optimization problem. His approach differs from ours along several dimensions. First, we do not consider transaction costs. Second, we start by assuming a specific utility function over wealth instead of specifying a loss function. Third, in our approach, the optimal target ratio is endogenously specified as part of the solution. Finally, we model the basis as an explicit state variable and solve for the trading policy in the presence of capital gains taxes in the multi-asset case.

Our approach borrows from the traditional transaction costs literature. Magill and Constantinides (1976), Taksar, Klass and Assaf (1988), Davis and Norman (1990) Dixit (1991), Dumas (1991) and and Shreve and Soner (1994) provide a characterization of the solution of the portfolio problem with a single risky asset and with proportional transaction costs. Akian, Menaldi and Sulem (1996) extend Davis and Norman's (1990) results to the multi-asset case. We use ideas from this literature to model a more complex and state-dependent form of transaction cost.

The rest of the paper proceeds as follows. In Section 3.2 we develop the model for the two-asset case. In Section 3.3 we numerically solve the two-asset model and analyze the properties of the optimal trading strategy, as well as the value of tax deferral and the intertemporal effect of capital gains taxes. In Section 3.4 we describe the portfolio problem of an investor facing the opportunity of investing in a mutual fund and in an index fund. In Section 3.5 we analyze the impact on introducing wash-sales restrictions on investment strategies. Section 3.6 concludes. Appendix D contains a brief description of the single asset problem that will be used to develop some intuition for the more general case. Proofs for all propositions are in Appendix A, Appendix F contains details of the numerical implementation of the model, while Appendix G contains tables and figures.

3.2 The Model

We consider the optimal portfolio choice problem of a risk averse investor with investment horizon of $T > 0$ periods. The agent derives utility only from consumption at the terminal period T and can trade in three kinds of assets available in the financial markets: two risky stocks (labeled "Stock 1" and "Stock 2") and a riskless money market instrument. The agent is subject to an income tax on dividends and a capital gains tax (rebate) on realized gains (and losses). In addition, the agent is subject to short-sale constraints on the risky assets. In this section, we provide the set-up for our analysis of this problem, lay down the notation and discuss our assumptions.

3.2.1 The Distribution of Asset Returns

The price processes of the two risky assets are taken as exogenously given and these generate the fundamental uncertainty in the model. The (ex-dividend) risky price processes are given by $\{S_i(t), i = 1, 2, t = 0, \ldots, T\}$. The cum-dividend returns on the two assets at date $t+1$ are given by $\tilde{R}_i(t+1) + \frac{D_i(t+1)}{S_i(t)}$ where $\tilde{R}_i(t+1) \equiv \frac{S_i(t+1)}{S_i(t)}$ denotes the ex-dividend return on asset i and $D_i(t+1)$ is the dividend received at $t+1$ from holding the i^{th} asset from date t to $t+1$. We assume that dividend yields d_i , $i = 1, 2$ on the two assets are constant, i.e., $D_i(t+1) = d_i S_i(t), i = 1, 2$. Moreover, the vector of ex-dividend returns $\tilde{R}_i(t)$ is independently and identically distributed over time with a finite state space.

Given the assumption of a finite state space on returns, the fundamental uncertainty in the model can be represented by an event-tree structure. That is, the uncertainty in the model is described by a probability space $\{\Omega, \{\mathcal{F}_t\}_{t=0}^{T}, P\}$ where Ω is the set of all possible paths of the risky assets and $\{\mathcal{F}_t\}_{t=0}^{T}$ denotes an increasing family of partitions of Ω generated by the price processes $\{S_i(t), i = 1, 2, t = 0, \ldots, T\}$. The probability measure P on Ω is induced by the probability distribution of returns. This set-up is standard. See, for example, Huang and Litzenberger (1988, Chapter 7). Recall that, since $\mathcal{F}_t$ is an increasing sequence of partitions, for a given node $a \in \mathcal{F}_t$ there exists a unique node $\mathbf{a}(a) \in \mathcal{F}_{t-1}$ from which a can be reached. The set $\mathbf{a}(a)$ is the (unique) set of *predecessors* of a.

3.2.2 Investment Strategies

At time $t = 0$ the investor inherits a portfolio $(X_1^e(0), X_2^e(0), Y^e(0))$ and cost bases of the risky assets, $(B_1^e(0), B_2^e(0))$ where $X_1^e(0) \geq 0$, $X_2^e(0) \geq 0$, $B_1^e(0) \geq 0$, $B_2^e(0) \geq 0$. The investor then initiates a dynamic investment strategy. In the course of implementing this strategy, the investor arrives at node a at time t with an amount $X_i(a, t)$, $i = 1, 2$ in stock i (*Market value of stock i*), and $Y(a,t)$ in a money market instrument ("the bank account"). During this period, a portfolio rebalancing takes place with $\Delta_i(a,t)$, $i = 1,2$ denoting the dollar *purchase* (if $\Delta_i(a,t) > 0$) or sale (if $\Delta_i(a,t) < 0$) of stock i at time t. The change in the bank account after this rebalancing (and after payment of taxes, if any, and receipt of dividends) is denoted by $\Delta_Y(a,t)$. Of course, if $\Delta_Y(a,t)$ is negative, there is withdrawal of funds from the bank account.

During period t, dividends are received on the risky assets which are subject to a tax rate τ_d. By assumption, there is no interest received on the bank account. Any sale of risky assets with embedded capital gains is subject to a capital gains tax of $\tau_g(t)$. The tax rate on capital gains is assumed to be constant and equal to τ_g for $t = 0, \ldots, T-1$ and is equal to zero at time T, i.e.,

$$\tau_g(t) = \begin{cases} \tau_d & \text{if} \quad t = 0, 1, \ldots, T-1 \\ 0 & \text{if} \quad t = T \end{cases} . \tag{3.1}$$

This is done to capture a feature of the US tax code that provides capital gains tax forgiveness at death.[2] A realized capital loss generates a tax rebate at the same rate.[3]

Immediately after rebalancing at time t, the position in the stocks and in the bond are:

$$X_i'(a,t) = X_i(a,t) + \Delta_i(a,t), \quad i = 1,2 \tag{3.2}$$

$$Y'(a,t) = Y(a,t) + \Delta_Y(a,t) \tag{3.3}$$

As the above description shows, an investment strategy prescribes the amount invested in the assets at every possible node and at all time periods. Formally, the investment strategy is an element of $\mathbb{R}^N$ where $N = \sum_{t=0}^{T-1} N_t$ and N_t is the number of elements in $\mathcal{F}_t$. We denote an investment strategy by $\delta_T \in \mathbb{R}^N$ where

$$\delta_T \equiv \{\Delta(a,t) \equiv (\Delta_1(a,t), \Delta_2(a,t), \Delta_Y(a,t)), \quad a \in \mathcal{F}_t, \quad t = 0, \ldots, T-1\}. \tag{3.4}$$

Next, we describe the treatment of capital gains and the evolution of investment holdings and cost bases in our model.

[2]The Canadian Tax code does not provides tax forgiveness at death. In this case we would have $\tau_g(T) = \tau_g$ also.

[3]We ignore some of the complexities of tax law such as off-setting requirements and an asymmetry in the treatment capital gains and losses. Moreover, we do not differentiate between long-term and short-term capital gains.

3.2.3 Treatment of Capital Gains

Our model of capital gains draws from Dammon, Spatt and Zhang (2001). We denote by $G_i(a,t)$ the capital gain/loss on asset $i = 1,2$ arising after rebalancing in node $a \in \mathcal{F}_t$ at time t. In order to calculate such gains/losses, it is necessary to define a "cost basis". The basis at every time is represented by the dollar cost of stocks held at time t (i.e., the *Book Value of shares*). We denote by $B_i(t)$ the the basis for stock i, $i = 1,2$ at time t. In what follows, we allow for *wash sales* for assets with embedded capital losses. That is, the investor is allowed to sell and immediately repurchase an asset for the purpose of realizing capital losses and resetting the basis. In Section 3.5 we deal with the case in which wash-sales are not allowed.[4] Ideally, if it was allowed in the tax code, the investor would like to track the costs of each group of purchased shares, and then choose what to sell according to the embedded gains/losses. However, this problem is path-dependent and it quickly becomes intractable to calculate optimal strategies.[5] Instead, we assume the investor elects to use the average cost method for calculating cost basis, and we model the basis as a weighted average of purchase prices.

Consider first an asset that has an embedded capital gain in (a,t), $a \in \mathcal{F}_t$, i.e., an asset i for which $B_i(a,t) < X_i(a,t)$. Suppose that the investor desires to make an additional investment of $\Delta_i(a,t)$ in this asset. Then, if $\Delta_i(a,t) > 0$ (i.e., if the investor wishes to buy more of this asset), no capital gains, and hence no capital gains taxes, are incurred. On the other hand, if $\Delta_i(a,t) < 0$ (i.e., if the investor is selling the asset), the investor realizes a capital gain equal to

$$-\Delta_i(a,t)\left(1 - \frac{B_i(a,t)}{X_i(a,t)}\right).$$

Suppose now that the asset has an embedded capital loss (i.e., $B_i(a,t) > X_i(a,t)$). Since wash sales are allowed, the investor will sell off the entire position in the risky asset to take advantage of the tax rebate and reset the basis before rebalancing his portfolio. Therefore, in this case the investor realizes a capital gain of $X_i(a,t) - B_i(a,t) < 0$, i.e., he will incur a capital loss realizing a tax rebate.

[4]According to the US tax code it is not possible to deduct a loss from a sale of a security if the same security or "substantially identical" securities are purchased within a window of 60 days centered around the date of the sale.

[5]Dybvig and Koo (1996) solve an intertemporal portfolio problem with capital gains taxes keeping track of every asset purchase in the determination of the correct basis. The complexity of the problem allows them to solve only a two-period, one-asset problem.

Thus, the realized capital gains $G_i(a,t), i = 1,2$ in node $a \in \mathcal{F}_t$ at date t are given by

$$G_i(a,t) = \begin{cases} -\Delta_i(t)\left(1 - \frac{B_i(a,t)}{X_i(a,t)}\right)\chi_{\{\Delta_i(a,t)<0\}} & \text{if } \; B_i(a,t) < X_i(a,t) \\ X_i(a,t) - B_i(a,t) & \text{if } \; B_i(a,t) \geq X_i(a,t) \end{cases}, \quad i = 1,2 \tag{3.5}$$

where χ is an indicator function equal to one when asset i is sold ($\Delta_i(a,t) < 0$) and zero otherwise.

3.2.4 The evolution of the cost bases and asset holdings over time

We now state the dynamics of the cost bases of the risky assets and the dollar holdings of various assets which form the state variables of the investor's problem.

Consider first the dynamics of asset holdings. Recall that, given a node $a \in \mathcal{F}_{t+1}$ the set $\mathbf{a}(a) \in \mathcal{F}_t$ represents the unique element of the filtration $\mathcal{F}_t$ from which a is reachable. Explicitly accounting for the information structure, we can write the ex-dividend return on stock $i = 1,2$ between time t and $t+1$ as follows:

$$\tilde{R}_i(a,t+1) = \frac{S_i(a,t+1)}{S_i(\mathbf{a}(a),t)}, \;\; a \in \mathcal{F}_{t+1}, \; \mathbf{a}(a) \in \mathcal{F}_t,$$

where $S_i(a,t+1)$ and $S_i(\mathbf{a}(a),t)$ denote the ex-dividend prices of asset $i = 1,2$ at time t, node $\mathbf{a}(a)$ and time $t+1$, node a respectively. The interest rate on the bank account is set to zero.

The dynamics of the investor's holdings in the risky stock and in the bank account for $t = 0,\dots,T-1$ and $a \in \mathcal{F}_{t+1}$ are given by

$$X_i(a,t+1) \;\; = \;\; (X_i(\mathbf{a}(a),t) + \Delta_i(\mathbf{a}(a),t))\tilde{R}_i(a,t+1), \;\; i = 1,2 \tag{3.6}$$

$$Y(a,t+1) \;\; = \;\; Y(\mathbf{a}(a),t) + \Delta_Y(\mathbf{a}(a),t) \tag{3.7}$$

Our assumption about the basis used to compute capital gains/losses is evident in equation (3.5) where we use the book-to-market factor $B_i(a,t)/X_i(a,t)$ to calculate the embedded gain in a sale. In other words, we assume that each basis evolves as a *moving average* of purchase prices. This means that a purchase increases the basis by the amount of the purchase itself while a sale decreases the basis by the average cost of purchased shares. For example, if $X_i(a,t) = \$15,000$, $B_i(a,t) = \$10,000$ and we purchase an amount $\Delta_1(a,t) = \$1,000$, then the basis will rise to $\$11,000$. On the other side, if we sell an amount $\$1,000$ it is assumed that the cost of the sold shares is an average of the purchase price, i.e., 10,000/15,000=0.667. Therefore the basis decreases by $1,000 \times 0.667 = \$667$ becoming $\$9,333$. Thus, if in state $\mathbf{a}(a) \in \mathcal{F}_t$ asset i has an embedded gain (i.e., $B_i(\mathbf{a}(a),t) < X_i(\mathbf{a}(a),t)$) and the investor sells the asset (i.e., $\Delta_i(\mathbf{a}(a),t) < 0$) the basis of this asset at time $t+1$, node $a \in \mathcal{F}_{t+1}$ is

$$B_i(a,t+1) = B_i(\mathbf{a}(a),t) + \Delta_i(\mathbf{a}(a),t)\left(\frac{B_i(\mathbf{a}(a),t)}{X_i(\mathbf{a}(a),t)}\right).$$

If the investor purchases more of asset i (i.e., $\Delta_i(\mathbf{a}(a),t) > 0$), then

$$B_i(a,t+1) = B_i(\mathbf{a}(a),t) + \Delta_i(\mathbf{a}(a),t).$$

If, however, asset i has an embedded loss (i.e., $B_i(\mathbf{a}(a),t) > X_i(\mathbf{a}(a),t)$), given that wash sales are allowed, the investor sells and immediately repurchases the asset with a loss to generate an immediate tax rebate. This resets the basis of the initial holdings to $B_i(\mathbf{a}(a),t) = X_i(\mathbf{a}(a),t)$. After the wash-sale the portfolio rebalancing takes place by an amount of $\Delta_i(\mathbf{a}(a),t)$. Therefore the basis $B_i(\mathbf{a}(a),t+1)$ after trading at time t will be equal to the value of asset i held after trade at time t, i.e.,

$$B_i(\mathbf{a}(a),t+1) = X_i(\mathbf{a}(a),t) + \Delta_i(\mathbf{a}(a),t).$$

To summarize, in response to a portfolio choice $\Delta_i(\mathbf{a}(a),t)$ at time t in state $\mathbf{a}(a) \in \mathcal{F}_t$, the dynamics of the bases $B_i(\mathbf{a}(a),t)$, $i = 1,2$ are as follows:

$$B_i(a,t+1) = \begin{cases} B_i(\mathbf{a}(a),t) + \Delta_i(\mathbf{a}(a),t)\left(1 + \chi_{\{\Delta_i(\mathbf{a}(a),t)<0\}}\left(\frac{B_i(\mathbf{a}(a),t)}{X_i(\mathbf{a}(a),t)} - 1\right)\right) & \text{if } \frac{B_i(\mathbf{a}(a),t)}{X_i(\mathbf{a}(a),t)} < 1 \\ X(\mathbf{a}(a),t) + \Delta_i(\mathbf{a}(a),t) & \text{if } \frac{B_i(\mathbf{a}(a),t)}{X_i(\mathbf{a}(a),t)} \geq 1 \end{cases}. \tag{3.8}$$

The description of the dynamics of the state variables of the investor's problem in response to an arbitrary investment strategy is now complete. We are now in a position to define the set of feasible investment strategies.

3.2.5 Feasible Investment Strategies

Let $\mathbf{B}(a,t) = (B_1(a,t), B_2(a,t))$ be a vector of bases at time t, node $a \in \mathcal{F}_t$. The following set describes the *solvency region* for the investor at time t in state $a \in \mathcal{F}_t$

$$\mathcal{S}_{\mathbf{B}(a,t)} = \left\{(x_1,x_2,y) \in \mathbb{R}^2_+ \times \mathbb{R} : y + \sum_{i=1}^{2}(x_i - \tau_g(t)(x_i - B_i(a,t))) \geq 0\right\} \tag{3.9}$$

The solvency region contains the portfolio positions that guarantee a non-negative *liquidating* value. The solvency region for our problem as the set of positions (X_1, X_2, Y) from which the agent can move to a position of non-negative investment in all the assets. From every point of the solvency region, the investor can sell all the asset and incur all taxes on gains an losses without facing personal bankruptcy.

There are three essential features of feasible investment strategies: (i) they are self-financing (ii) short-selling is not allowed and (iii) they insure a non-negative after-tax (liquidating) wealth at all times t and all possible nodes $a \in \mathcal{F}_t$. Moreover, the set of feasible investment strategies depends on the initially inherited portfolio position and bases, $\mathbf{Z}^e(0) \equiv [X_1^e(0), X_2^e(0), Y^e(0), B_1^e(0), B_2^e(0)]$. We denote by $\Phi_T(\mathbf{Z}^e(0)) \subset \mathbb{R}^N$ the set of feasible trading strategies starting from the inherited position $\mathbf{Z}^e(0)$. The following definition describes such set.

Definition 3.2.1 *Let* $\mathbf{Z}^e(0)$ *be such that* $(X_1^e(0), X_2^e(0), Y^e(0)) \in \mathcal{S}_{\mathbf{B}^e(0)}$. *An investment strategy* δ_T *(as defined in (3.4)) is* feasible *for* $\mathbf{Z}^e(0)$*, i.e.,* $\delta_T \in \Phi_T(\mathbf{Z}^e(0))$*, if, for every* $a \in \mathcal{F}_t$ *and* $t = 0, \ldots, T$

1. $Y'(a,t) \leq Y(a,t) - \sum_{i=1}^{2} \Delta_i(a,t) - \tau_g(t) \sum_{i=1}^{2} G_i(a,t) + (1-\tau_d) \sum_{i=1}^{2} d_i(X_i(a,t) + \Delta_i(a,t))$ *(Self-financing)*
2. $\Delta_i(a,t) \geq -X_i(a,t)$, $i = 1, 2$. *(No short-selling)*
3. $(X_1(a,t), X_2(a,t), Y(a,t)) \in \mathcal{S}_{\mathbf{B}(a,t)}$. *(Solvency)*

In the above, $X_i(0) = X_i^e(0)$, $i = 1, 2$, $Y(0) = Y^e(0)$ *and the evolutions of* $X_i(a,t)$, $Y(a,t)$, $B_i(a,t)$, $i = 1, 2$ *are given by (3.6), (3.7) and (3.8).*

The self financing constraint is the budget constraint of the investor. It says that the bank balance of the investor after the rebalancing at time t is no greater than the initial bank balance plus the dividend received at time t (net of taxes) less the amount spent on the new investment less the payment of taxes on realized gains, if any. The constraint on short-selling is meant to capture the difficulties in the real world in going short in a risky asset, especially for an individual investor. Costantinides (1983) shows that if securities can be sold short with full use of the short-sale proceeds and if wash sales are allowed (i.e., if it is possible to sell and immediately repurchase an asset for the purpose of realizing capital losses and resetting the basis), then optimal investment and liquidation decision are separable from the consumption decision. Therefore, if short-selling is allowed, the optimal trading strategy is to realize losses as soon as they arise and defer gains indefinitely. In the real world, not only are there restriction on wash sales, but short selling is costly. Typically, one does not get the full use of the proceeds of short-sales and must additionally satisfy a margin requirement to cover any losses arising from sudden price increases.

The non-negative liquidating wealth constraint ensures that the set of feasible investment strategies is non-empty. It is always possible, in fact, to truncate any investment strategy at the point in which it violates the non-negativity constraint, liquidate everything, invest the (non-negative) amount in the riskless asset and then do nothing from there on. Moreover, the non-negativity constraint implicitly imposes a bound on the amount that can be borrowed in each period. In the absence of such a constraint, depending on the penalty for negative wealth, an investor may want to borrow an unlimited amount. When there is an infinite number of trading opportunities, this constraint is one way of preventing arbitrage opportunities. See Dybvig and Huang (1988) and Back and Pliska (1991).

In Appendix D we provide an intuitive characterization of the feasibility set for the one-asset one-period problem. The example shows that the budget set when an asset has an embedded gain contains a "kink" around the no-trade point. As in the transaction costs case, this kink is responsible for the no-trade result we will extensively document in Section 3.3 and the consequent loss of diversification induced by the existence of capital gains taxes. Moreover, we also note that if we do not impose non-negative liquidating wealth, the feasibility set might be empty and the problem would be ill-defined. The following proposition, proved in Appendix A, formally characterizes the set of feasible strategies for our problem

Proposition 3.2.2 *The feasibility set* $\Phi_T(\mathbf{Z}(0))$ *is non-empty and compact.*

We now describe the preferences of the individual and the optimization problem he faces. In what follows, unless necessary, we suppress the dependence of investment strategies and state variables on the nodes a of the event tree.

3.2.6 Preferences

The investor seeks to maximize the expected utility of terminal consumption. Preferences are assumed to be described by a time-separable, state independent isoelastic utility function

$$U(w) = \frac{1}{1-\gamma} w^{1-\gamma}, \quad w \geq 0, \gamma > 0,$$

where γ is the risk aversion coefficient. There is no intermediate consumption. At time $t = 0$ the investor inherits a portfolio $(X_1(0), X_2(0), Y(0))$ and cost bases $(B_1(0), B_2(0))$ for the risky assets. Let

$$\mathbf{Z}(0) = [X_1(0), X_2(0), Y(0), B_1(0), B_2(0)]$$

represent the initial state variables at time $t = 0$. The investor chooses an investment strategy $\delta_T \in \Phi_T$ to maximize the expected utility of terminal wealth.

Using the set-up stated above, the investor's (static) maximization problem can be explicitly formulated as follows

$$J_0(\mathbf{Z}(0)) = \max_{\delta_T \in \Phi_T(\mathbf{Z}(0))} E_0^P \left[\left(\sum_{i=1}^{2} X_i(T) + Y(T) \right)^{1-\gamma} \right] \tag{3.10}$$

where $X_i(T)$ and $Y(T)$ are attainable through a feasible policy δ. The boundary condition reflects the "tax-forgiveness" at death of the US tax code.[6]

3.2.7 Dynamic Programming Formulation

Since by Proposition 3.2.2 the feasibility set is compact and the objective function is continuous, the problem is well defined and admits a solution. This allows us to rewrite the problem as a dynamic programming problem by invoking the Bellman's principle of optimality. Let β be the subjective discount factor for utility, then the investor's indirect utility function $J_t(\cdot)$ can be expressed as the solution of the following dynamic programming problem for $t = 0, \ldots, T-1$

$$J_t(\mathbf{Z}_t) = \max_{\mathbf{\Delta}_t \in \Phi(\mathbf{Z}_t)} E_t^P[\beta J_{t+1}(Z_{t+1}(\Delta_t))] \tag{3.11}$$

[6]In the Canadian code there is no tax forgiveness at death. In this case the boundary condition would be

$$U(\mathbf{Z}_T) = \frac{1}{1-\gamma} \left(\sum_{i=1}^{2} (X_i(T) - \tau(X_i(T) - B_i(T)) + Y(T) \right)^{1-\gamma}.$$

From now on we adopt the US code of tax forgiveness at death as a terminal condition.

where $\Phi(\mathbf{Z}_t)$ denotes the set of feasible trades $\mathbf{\Delta}_t = (\Delta_1(t), \Delta_2(t))$ available at time t from state $\mathbf{Z}_t$,[7] $\mathbf{Z}_{t+1}(\Delta_t)$ is given by (3.6), (3.7) and (3.8) and

$$J_T(\mathbf{Z}_T) = \frac{1}{1-\gamma}\left(\sum_{i=1}^{2} X_i(T) + Y(T)\right)^{1-\gamma}. \tag{3.12}$$

Since $J(\cdot)$ is homothetic we can normalize by the beginning of period (liquidating) wealth

$$W(t) = Y(t) + \sum_{i=1}^{2}(X_i(t) - \tau_g(t)(X_i(t) - B_i(t)))$$

and thus reduce the number of state variables. We indicate with $x_{i,t} = X_i(t)/W(t)$, $i = 1, 2$ the fraction of wealth allocated to stock $i = 1, 2$, $\theta_{i,t} = B_i(t)/X(t)$ the basis-to-price ratio of asset i and $\alpha_{i,t} = \Delta(t)/W(t)$ the fraction of wealth traded in stock i at time t. Since we do not consider consumption, the self financing requirement implies a natural relationship between the amount traded in the risky securities Δ_i and in the money market account Δ_Y, i.e.,

$$\Delta_Y(t) = -\sum_{i=1}^{2}(\Delta_i(t) + \tau(t)G_i(t)) + (1-\tau_d)\sum_{i=1}^{2}(X_i(t) + \Delta_i(t))d_i. \tag{3.13}$$

Therefore equation (3.7) can be rewritten as

$$Y(t+1) = Y(t) - \sum_{i=1}^{2}(\Delta_i(t) + \tau_g(t)G_i(t)) + (1-\tau_d)\sum_{i=1}^{2}(X_i(t) + \Delta_i(t))d_i. \tag{3.14}$$

The normalized investment in the money market account is

$$y_{1,t} = \frac{Y(t)}{W(t)} = \frac{W(t) - \sum_{i=1}^{2}(X_i(t) - \tau_g(X_i(t) - B_i(t)))}{W(t)} = 1 - \sum_{i=1}^{2}(x_{i,t} - \tau_g x_{i,t}(1-\theta_{i,t})) \tag{3.15}$$

and we can reduce also the space of control by simply considering the fraction of liquidating wealth $\alpha_{i,t}$ traded in asset i, $i = 1, 2$. Let

$$\mathbf{z}_t = [x_{1,t}, x_{2,t}, \theta_{1,t}, \theta_{2,t}]$$

be the vector of state variables for the normalized optimization problem. The dynamics of $\mathbf{z}_t$ are easily computed as follows[8]

$$x_{i,t+1} = (x_{i,t} + \alpha_{i,t})\tilde{R}_{i,t+1} \cdot \tilde{M}_{t+1}(\mathbf{z}_t; \alpha_{1,t}, \alpha_{2,t}), \quad i = 1, 2 \tag{3.16}$$

$$\theta_{i,t+1} = \begin{cases} \frac{\theta_{i,t}x_{i,t} + \alpha_{i,t}\left(1 + \chi_{\{\alpha_{i,t}<0\}}(\theta_{i,t}-1)\right)}{x_{i,t}+\alpha_{i,t}} \frac{1}{\tilde{R}_{i,t+1}} & \text{if } \theta_{i,t} < 1 \\ \frac{1}{\tilde{R}_{i,t+1}} & \text{if } \theta_{i,t} > 1 \end{cases}, \quad i = 1, 2, \tag{3.17}$$

[7]Note that $\Phi(\mathbf{Z}_t)$ is a set of *trades* while $\Phi_T(\mathbf{Z}(0))$ is a set of *strategies*.

[8]Note that $\theta_{i,t}$ can be written as $\frac{B_i(t)/W(t)}{X_i(t)/W(t)}$ for $i = 1, 2$.

where

$$\tilde{M}_{t+1}(\mathbf{z}_t;\alpha_{1,t},\alpha_{2,t}) = \frac{W(t)}{W(t+1)}$$

$$= \frac{W(t)}{Y(t+1)-\sum_{i=1}^{2}(X_i(t+1)-\tau_g(t+1)(X_i(t+1)-B_i(t+1)))}. \quad (3.18)$$

Using (3.1), (3.6), (3.8) and (3.14) we conclude that, for $t=0,1,\ldots,T-2$,

$$\tilde{M}_{t+1}(\cdot) = \left[1+(1-\tau_g(t))\sum_{i=1}^{2}\Big((x_{i,t}+\alpha_{i,t})(\tilde{R}_{i,t+1}-1)\Big)+(1-\tau_d)\sum_{i=1}^{2}d_i(x_{i,t}+\alpha_{i,t})\right]^{-1}, \quad (3.19)$$

and, for $t=T-1$,

$$\tilde{M}_T(\cdot) = \Bigg[1+\sum_{i=1}^{2}\Big((x_{i,T-1}+\alpha_{i,T-1})(\tilde{R}_{i,T}-1)+d_i(1-\tau_d)(x_{i,T-1}+\alpha_{i,T-1})\Big)$$
$$+\tau_g(T-1)\sum_{i=1}^{2}\chi_{\{\theta_{i,t}<1\}}\Big((1-\theta_{i,T-1})(x_i+\alpha_{i,T-1}\chi_{\{\alpha_i<0\}})\Big)\Bigg]^{-1} \quad (3.20)$$

We define the restriction of the set of feasible trades $\Phi(\mathbf{Z}(t))$ (see equation (3.11)) on the set of controls $\mathbf{\Delta}_t = (\Delta_1(t),\Delta_2(t))$ implied by the self-financing constraint (3.13) as follows

$$\Phi^Y(\mathbf{Z}_t) = \left\{\mathbf{\Delta}_t \in \mathbb{R}^2 : \left(\mathbf{\Delta}_t, -\sum_{i=1}^{2}(\Delta_i(t)+\tau(t)G_i(t))+(1-\tau_d)\sum_{i=1}^{2}(X_i(t)+\Delta_i(t))d_i\right)\in\Phi(\mathbf{Z}_t)\right\}.$$

This allows us to characterize the feasibility set from state $\mathbf{z}_t$ for the controls $\boldsymbol{\alpha}_t = (\alpha_{1,t},\alpha_{2,t})$ as

$$\phi(\mathbf{z}_t) = \{\boldsymbol{\alpha}_t \in \mathbb{R}^2 : \boldsymbol{\alpha}_t W(t) \in \Phi^Y(\mathbf{Z}_t)\}. \quad (3.21)$$

Finally, if we let

$$v_t(\mathbf{z}_t) = \frac{J_t(\mathbf{Z}_t)}{W(t)^{1-\gamma}}, t=0,1,\ldots,T.$$

and use the homotheticity of $J(\cdot)$, we can reformulate the dynamic programming problem (3.11) as follows

$$v_t(\mathbf{z}_t) = \max_{\boldsymbol{\alpha}_t\in\phi(\mathbf{z}_t)} E_t^P\left[\beta v_{t+1}(z_{t+1})\left(\frac{1}{\tilde{M}_{t+1}(\mathbf{z}_t;\alpha_{1,t},\alpha_{2,t})}\right)^{1-\gamma}\right] \quad (3.22)$$

s.t. (3.16), (3.17), (3.19), (3.20) and with boundary conditions

$$v_T(\mathbf{z}_t) = \frac{1}{1-\gamma}.$$

We assume that the vector of returns $(\tilde{R}_{1,t+1},\tilde{R}_{2,t+1})$ follows a bi-variate trinomial process as discussed in Appendix F. We discretize the state space $\mathbf{z}_t = \{x_{1,t},x_{2,t},\theta_{1,t},\theta_{2,t}\}$ into a four-dimensional grid, solve the problem at period T and then solve backwards recursively. Further details of the numerical implementation are provided in Appendix F.

3.3 Numerical Results for the Two-Asset Case

In this section we present the results from the solution of the optimal investment problem under capital gains taxes with two risky assets.

We calibrate our benchmark numerical solution by choosing parameters as follows. We assume that both risky assets have the same gross return, $\mu_1 = \mu_2 = 10\%$ and the same volatility $\sigma_1 = \sigma_2 = 30\%$. The dividend yields on both assets are set equal to zero ($d_1 = d_2 = 0$).[9] The correlation coefficient is set to $\rho = 0.5$. The risk aversion parameter for the investor is $\gamma = 5$ and the subjective discount factor is set to $\beta = .96$. The tax rate on realized gains and losses and on dividends is assumed to be $\tau_g = \tau_d = 25\%$. We solve a five-period model, so $T = 5$. In Table XII we report the parameter values used in the numerical solution of the benchmark two asset case.

3.3.1 Effect of Taxes on Optimal Diversification

We begin by analyzing the investment strategies over a wide variety of initial asset holdings and bases for the two assets. In particular, we examine how diversification motives and tax motives interact in determining the optimal trading strategy. We will compare this strategy with the strategy prevailing in the absence of taxes, which we refer to as the *Merton Problem*, as in Merton (1971).

Figure G.1 reports the trades in asset 1 over a wide area of the initial state space. In each panel, the holdings of asset two (x_2) and the gain or loss in asset 2 (θ_2) are held constant. Each line on the plot shows the optimal trade as the holding of asset 1 (x_1) is varied. One line shows the investment strategy for an agent endowed with a capital gain in the first asset ($\theta_1 = .2$). The second line shows the strategy for an embedded loss ($\theta_1 = 1.1$). Finally, the solid line shows that investment strategy for the no-tax case.

Panel A presents the situation where the investor has a low holding of asset 2, and asset 2 has an embedded gain ($x_2 = 0.1, \theta_2 = .2$). Panel B shows the same holding of asset 2, but we assume that the asset has an embedded loss ($x_2 = 0.1, \theta_2 = 1.1$). Similarly, Panels C and D, consider the case of an over-exposure of asset 2 ($x_2 = 0.7$) with an embedded gain ($\theta_2 = .2$) and loss ($\theta_2 = 1.1$) respectively.

[9]In a separate paper we plan to fully analyze the effect of dividends and capital gains taxes on investment strategies.

Our first observation about Figure G.1 is that, when asset 1 has an embedded gain ($\theta = 0.6$), there is a region of holdings where the investor chooses neither to buy, nor to sell. We refer to this flat part of the curve as a "no-trade" region. We can see from the Merton line, that the optimal strategy without taxes would be to sell the asset in this region. However, capital gains taxes act as a transaction cost on the sale of the asset with an embedded gain. Over a region of holdings in asset 1 where the Merton problem would imply a sale, this implicit transaction cost outweighs the cost of sub-optimal diversification. These costs due to taxes on realized capital gains generate a no-trade region exactly as in a traditional transaction costs problem. When the holding in asset 1 increases and the over-exposure in it becomes larger than the investor is willing to bare, some transaction costs in the form of capital gains taxes on realization are accepted in order to reduce the diversification costs. However, the sale is always for an amount less than the equivalent Merton problem sale.

The no-trade region can be exactly mapped to a no-transaction cone. As shown in Appendix D and Figure D.1, the feasibility set for an investor holding an asset with embedded gains has a kink at the no-trade point. As in the transaction cost literature, this non-linearity is responsible for the presence of a region where the investor chooses not to trade. Since capital gains taxes are state dependent, our model differs from the standard transaction-cost literature. We can see this by comparing the above results with the investment strategy for the agent with an embedded loss. In this case, we do not observe a no-trade region. More importantly, note that, just as in the Merton problem, this investor trades to a fixed holding of asset 1 regardless of the initial allocation of this asset. The feasibility set for this investor is linear in the asset with the loss—just as in the no-tax case—but is shifted upwards by the amount of the tax rebate. The intuition is again contained in Appendix D. This result holds for an investor with any initially endowed embedded loss. The reason for this is that, by performing a wash sale, the investor with a capital loss will always reset the basis (i.e., $\theta_1 = 1$ after a wash sale) and rebalance based on this neutralized position. Comparing the capital-loss investments to the no-tax investments, we see that the agent holds a larger constant portfolio in the asset with the capital loss. The investor is willing to hold more of the taxable capital-loss asset since this asset embeds a tax-timing effect which allows the immediate realization of losses and (possibly infinite) deferral of capital gains.

To further examine the effect of capital gains taxes on optimal diversification, we graphically show the region of no-trades as a function of the holdings of the two risky assets in Figure G.2. In the black region, the investor chooses not to trade in either risky asset. The grey region shows positions where only one risky asset is traded, and in the white region both assets are traded. The intersection of the horizontal and vertical black lines at $(0.2, 0.2)$ represents the *degenerate* no-trade region in the no-tax (Merton) case. Since the grey and black regions show conditions in which the investor avoids trading for tax reasons, the size of the white region in each Panel gives an idea of the degree of loss of diversification due to capital gains taxes. Not surprisingly, we note that in Panel A, where both assets have large embedded gains, the black and grey region covers most of the state space. Investors with very low holdings will always trade because taxes place no constraint on purchases. In addition, investors with very high holdings will trade due to increasing diversification costs.

An intuitive way of summarizing the above results is the following:

Result 1 (Trades in two assets) *When wash sales are allowed, the trading strategy is characterized by (i) a "No-trade region" if the asset has an embedded gain, (ii) a constant after-rebalancing holding of the asset with an embedded loss which exceeds the no-tax holding.*

The above plots refers to the benchmark case in which $\rho = 0.5$. We know that in this case, the two assets are identically distributed (see Appendix F), and therefore the patterns of trades for asset 2 will be exactly identical. Comparing the four panels in figure G.1 which show the investment strategy for asset 1 as we change the holdings and tax status of asset 2, we see that Panels B and D are exactly identical. This is due to the aforementioned fact that, with wash sales, any position in asset 2 which embeds a loss is equivalent to a neutralized position. The after-rebalancing portfolio holding of asset two is therefore identical for any initial holding and tax status, provided that there is an embedded loss. Comparing Panels A and C, we see a very slight difference in the strategies, particularly around the no-trade region. Here, the investor is unwilling to fully rebalance his asset 2 holdings because of the embedded gain in asset 2. Therefore, as the initial holdings of asset 2 change, the agent changes his investment in asset 1 to compensate for the locked-in holding in asset 2. In addition, we see in Figure G.2 that the no-trade regions are, in general, not rectangular. This reinforces the existence of cross-effects of the holding of one asset on the strategy in the other asset. Next, we examine this effect in more detail by explicitly analyzing, for different values of the correlation among assets, how the investment in one asset is affected by the holding and tax status of the other asset.

Figure G.3 shows the trades in both assets as a function of holdings in asset 1 when (i) asset 1 has an embedded gain ($\theta_1 = .2$), (ii) asset 2 is neutralized ($\theta_2 = 1$), and (iii) the investor is under-exposed in asset 2 ($x_2 = 0.1$), compared to the Merton case. Panel A reports the case in which assets are negatively correlated ($\rho = -0.5$) while Panel B refers to the case of positively correlated assets ($\rho = 0.9$). The Merton (no-tax) portfolio holdings in the two assets are ($x_1 = 0.263, x_2 = 0.223$) in Panel A and ($x_1 = 0.076, x_2 = 0.208$) in Panel B.

In this figure, we see that the friction created by the capital gains taxes on asset 1 causes the investor to choose to sell less of this asset than diversification motives would suggest. As $x1$ increases across the no-trade region (dashed line), the investor becomes more and more exposed to asset 1. When asset 2 is negatively correlated with asset 1, as in Panel A, the agent compensates for the over-exposure to asset 1 by purchasing more of asset 2, which acts as a *complement.* On the other side, if asset 2 is strongly positively correlated with asset 1, as in Panel B, the agent divests himself of asset 2 as the over-exposure to asset 1 increases. Since the two assets are essentially the same in this case, a sale of asset 2 can perfectly substitute for the inability to sell asset 1.

This can be summarized in the following result.

Result 2 (Cross-effects of Asset Holdings) *The quantity held in one asset can influence the investment strategy in the other asset. Whenever capital gains taxes prevent optimal diversification in one asset, the investment strategy in the other asset will change to compensate for this friction. The direction and degree of the change depends on whether assets are complements (negatively correlated) or substitutes (positively correlated).*

3.3.2 Value of Optimal Exercise of the Tax-Timing Option

In this section, we analyze the value function for the investor trading in two risky assets. In order to gain intuition about the effect of taxation on the investor's utility, we compare the value function obtained from the two-asset taxable portfolio problem with the no-tax case. We the compare the two utility functions by calculating the percentage of extra wealth necessary to make the holder of the non-taxable assets indifferent to holding assets subject to taxation. In the following we will refer to this measure as to the *certainty equivalent* cost of the tax-on realization case versus the non-tax case.[10] Values greater than zero indicate states where investing in taxable assets is preferred. We interpret this quantity as the value of optimal tax deferral.

In Figure G.4, we report this value for a variety of different values of the state variables. In each panel, we plot the value as a function of the holding and basis of asset 1, keeping the holding and basis of asset 2 constant. The highest value for the taxable case in each panel is achieved when the investor has a high holding of asset 1 and asset 1 has a large embedded gain (north-west corner). There are two things that may seem puzzling in Figure G.4. First, the taxable investor is always better off than the non-taxed investor. Second, the larger the tax liability, the better off is the taxable investor relative to the non-taxed investor.

An easy way to examine the intuition behind the first puzzle is to consider the case of assets that, rather than being subject to taxes only when they are sold, instead realize taxes each period based on the accrued return. Solving for the optimal trading strategy in this case is equivalent to solving a no-tax problem where the returns are scaled by a factor $1 - \tau_g$. This transformation leaves the Sharpe ratio μ/σ unchanged, and consequently, the value function is identical to the no-tax case for an investor with power utility.[11] Intuitively, taxation upon accrual is equivalent to paying for an *insurance policy* from the government which reduces both the mean and variance of returns. Therefore, symmetric taxation on accrual neither increases nor decreases the utility of the investor. Capital gains taxes, however, are paid upon realization and therefore allow the investor substantially more flexibility in timing the tax payments over the investment horizon. Capital losses are realized immediately through wash sales. Capital gains can be delayed and the investor may even avoid these taxes entirely because of tax forgiveness at death. Combining the insurance effect and the ability to defer gains, the investor paying tax on realization is overall better off than the investor paying tax on accrual.

The second puzzle is a consequence of the fact that we normalize the state space by the liquidating wealth. Consider an investor with a capital gain, and an investor with a capital loss. Figure G.4 shows that the gain investor is better off. By construction, both investors have the same liquidating wealth. However, the liquidating wealth for the investor with the gain includes a large tax liability. This liability can be deferred, possibly indefinitely. This option to defer capital gains

[10] In welfare economics, this measure is known as the Kaldor-Hicks compensation measure.

[11] This result can be proved by looking at the value function emerging from the solution of a continuous time portfolio selection problem. See, for example, Merton (1971).

taxes adds value to the portfolio of the gain investor and results in a higher utility level. Using similar intuition, when we compare the four panels, we see that utility is highest in Panels A and C where the second asset has a large embedded gain.

Finally, observe that the utility level is constant in the loss region ($\theta_1 \geq 1$). The reason for this is the possibility of wash sales which, as we commented above, lead to identical portfolio choices from every loss position. The quantity of the tax rebate gained is different for every loss position. Normalizing by the liquidating wealth absorbs this effect and leads to a flat surface in the loss region.

In Table XIII we report the value of the tax-deferral option as a function of the volatility in the two risky asset and of the coefficient of relative risk aversion γ. We recall that the value of the tax-deferral option is defined in terms of the extra percentage wealth that need to be given to the investor paying taxes on accrual in order make him as well off as being able to pay taxes on realization. Notice that, from Table XIII, the value of the tax deferral option is *decreasing* in the volatility of the assets and in the risk aversion of the investor. The reason for this can be explained with an analogy to option pricing theory if we interpret the volatility of the assets and the risk aversion as component of the *exercise price* of the deferral option. The option to defer, in fact, is in the money when there is an embedded gain and is gradually more and more valuable the higher the embedded gain. If the investor decide to exercise this option (i.e., if he decides to defer the embedded gain) he will have to pay a price in terms of "lost diversification". Such a "diversification price" is higher (i.e. the option is less valuable) the higher is the volatility and the higher is the risk aversion because the investor, by exercising the deferral option is implicitly accepting to hold a larger fraction of an increasingly risky asset. In an intertemporal problem the tax deferral option is actually a *compound* option. In this context, the moneyness tomorrow depends on the basis and the asset holding today, as well as the volatility, which affects the basis and asset holding tomorrow. Risk aversion affects the payoff that we get in the future by exercising the option.

In Table XIV we report the sensitivity of the value of tax-deferral option to changes in the correlation between the risky assets. This value decreases as the correlation increases. Again we can provide an intuition by borrowing from option pricing theory. By investing in two securities, the investor is actually holding a portfolio of tax deferral options. When the underlying assets are highly correlated, holding such a portfolio of option is less valuable since the two options tend to be in the money or out of the money at the same time. Therefore, for highly correlated assets, the portfolio of tax deferral option resembles more to an option on a portfolio which, of course, is less valuable than a portfolio of options. On the contrary, when assets are negatively correlated, the investor can fully exploit the benefit of optimally and independently exercising the tax deferral option separately on the underlying assets.

We summarize our results for this section in the following:

Result 3 (Value of Tax Deferral) *The value of optimal tax deferral (i) increases in the asset holdings for assets with embedded gains, (ii) is constant in the asset holdings for assets with embedded losses, (iii) decreases in the volatility of the assets, (iv) decreases in the risk aversion of the investor and (v) decreases in the correlation of the assets.*

3.4 Effect of Capital Gains Taxes on Fund Trading

It is well-known that, without taxes and within a broad class of preferences (see, for example, Cass and Stiglitz (1970)) two-fund monetary separation holds, i.e., trading in two (or more) securities and a riskless asset is equivalent to trading in the riskless asset and in a fund that combines the two securities in a fixed proportion. The following proposition formalizes this concept in the context of our model.

Proposition 3.4.1 *In the absence of capital gains taxes, there always is a portfolio of the two asset such that trading in this fund is equivalent to trading in the two assets separately (i.e. two-fund separation holds). Moreover, the return $\tilde{R}_m$ on the fund is given by*

$$\tilde{R}_m = \omega \tilde{R}_1 + (1-\omega)\tilde{R}_2 \tag{3.23}$$

where

$$\omega = \frac{x_1 + \alpha_1^*(x_1,x_2)}{x_1 + \alpha_1^*(x_1,x_2) + x_2 + \alpha_2^*(x_1,x_2)}. \tag{3.24}$$

and $(\alpha_1^*(x_1,x_2), \alpha_2^*(x_1,x_2))$ are the optimal trades in absence of taxes.

In the presence of capital gains taxes, two-fund monetary separation does not hold: trading in multiple securities separately is different than trading in a fund. The nature of this difference depends heavily on the tax treatment of capital gains and losses within a fund.

In this section, we investigate the balance between tax-timing flexibility and diversification by considering two classes of securities, each of which, as we will make clear later, can be thought of as a form of fund characterized by a specific tax treatment of capital gains taxes. The first class of securities we analyze closely resembles most of the common *open-ended mutual funds.* In this category, shareholders are forced to pay taxes on the capital gains realized and distributed by the fund manager. In particular, every trade that the mutual fund manager makes to keep the fund in the desired proportions will cause the fund holder to incur capital gains or losses which are taxed. Even though the mutual fund holder does not directly control these trades, she is still forced to pay taxes based on them.

The second class of mutual funds is intended to capture the tax properties of some tax-efficient funds such as *index funds* and exchange-traded funds (ETF). The exact tax treatment of these funds is complicated and beyond the scope of this work. Instead, we simply model an important feature of this type of fund: the fund minimizes tax incurring on trades within the fund. We model this feature by assuming that trades within the fund do not generate tax liabilities or benefits to the fund holder. The index fund holder keeps track of the combined weighted cost basis of the fund, and pays taxes only on fund redemption based on the combined basis.

We will use the term *mutual fund* to refer to instrument in the first class and, with a slight abuse of terminology, *index funds* to refer to instruments in the second class. The comparison among the portfolio problem with (i) two separate assets, (ii) a mutual fund investing in the two assets and (iii) an index fund composed of the two constituent assets, provides a natural and interesting way of analyzing the interaction between optimal diversification and optimal tax-realization. In what follows we formalize the portfolio problems for these two classes of securities, numerically solve the allocation problem and compare results with the two-asset case.

3.4.1 Capital Gains Taxes and Mutual Funds

Let us consider the same economy as in Section 3.2 with the difference that now the investor has to allocate her wealth between a riskless asset and a mutual fund. We assume that the mutual fund manager invests in the two risky securities with the purpose of maintaining a *fixed* proportion between the two risky securities. The sole objective of the mutual fund manager is to trade to maintain the desired balance between risky assets in the fund.[12] By investing in a mutual fund, the investor is purchasing a security that guarantees her a pre-determined relative allocation between the two risky securities. Here, we assume that all the realized gains and losses generated by the mutual funds are distributed to the the investor at the end of every period.[13] In other words, the investor is forced to realize gains on trades that she delegated to the mutual fund manager.

Let α_m be the fraction of wealth invested in the mutual fund and ξ be the ratio of the two risky assets which the fund manager has promised to maintain.[14] Given a portfolio decision α_m and a portfolio ratio ξ we can compute the *implied* individual trades in asset 1 and 2 that the manager must perform. These are the fraction of wealth that are actually traded (*turnover*) in the individual assets by a fund manager who follows an allocation rule characterized by the ratio ξ. Starting from a state (x_1, x_2), the manager trades in order to achieve the desired balance between assets $((\alpha_1^m, \alpha_2^m))$ can be easily computed by solving the following system of equations

$$\frac{x_1 + \alpha_1^m}{x_2 + \alpha_2^m} = \xi \tag{3.25}$$

$$\alpha_1^m + \alpha_2^m = \alpha_m \tag{3.26}$$

Using the notation in Section 3.2, the intertemporal portfolio problem that the investor is solving can therefore be written as follows

$$v_t^F(\mathbf{z}_t) = \max_{\alpha_m \in \phi_m(\mathbf{z}_t)} E\left[\frac{1}{1-\gamma}\beta v_{t+1}^F(\mathbf{z}_{t+1}(\alpha_{m,t}))\left(\frac{1}{M_{t+1}(\mathbf{z}_t;\alpha_{m,t})}\right)^{1-\gamma}\right] \tag{3.27}$$

where $\phi_m(\mathbf{z}_t)$ is the one-dimensional equivalent of the feasibility set $\phi(\mathbf{z}_t)$ in (3.21) and where the state variables $\mathbf{z}_t = (x_{1,t}, x_{2,t}, \theta_{1,t}, \theta_{2,t})$ and $\tilde{M}(\mathbf{z}_t; \alpha_{m,t})$ evolve as in (3.16), (3.17), (3.19) and (3.20) when α_i is replaced by α_i^m , $i = 1, 2$. We noticed above that the ratio ξ is the *after trade* ratio in the two securities. Let

$$\omega_1 = \frac{x_1 + \alpha_1^m}{x_1 + x_2 + \alpha_m}, \quad \omega_2 = 1 - \omega_1$$

be the fraction of wealth invested in asset 1 after a trade. By (3.26), we can write

$$\xi = \frac{x_1 + \alpha_1^m}{x_2 + \alpha_2^m} = \frac{\omega_1}{1 - \omega_1}.$$

[12]The exact form of the mutual fund manager's objective function is not important. What is important here is to give the fund manager a motive for trade, i.e., a reason to generate high turnover, as it is often observed within traditional open-ended fund.

[13]To avoid corporate taxation, mutual funds are required to distribute a minimum of 90% of their income and capital gains to shareholders.

[14]If for example the desired ratio of asset 1 to asset 2 is 60:40, then $\xi = 1.5$.

The above shows that there is a one-to-one correspondence between the ratio that a fund manager guarantees and the fraction of wealth in each individual security that this ratio implies. This allows us to re-express the problem more directly as a constrained optimization problem in the variables α_1^m and α_2^m as follows

$$v_t^F(\mathbf{z}_t) = \max_{(\alpha_1^m,\alpha_2^m)\in\phi(\mathbf{z}_t)} E\left[\frac{1}{1-\gamma}\beta v_{t+1}^F(\mathbf{z}_{t+1}(\alpha_1^m,\alpha_2^m))\left(\frac{1}{M_{t+1}(\mathbf{z}_t;\alpha_1^m,\alpha_2^m)}\right)^{1-\gamma}\right] \tag{3.28}$$

s.t.

$$\frac{x_1+\alpha_i^m}{x_1+x_2+\alpha_m} = \omega_i, \quad i=1,2. \tag{3.29}$$

and (3.16), (3.17), (3.19) and (3.20).[15]

Given that this problem is a constrained two asset problem, it is straightforward to show that the value function obtained by solving the mutual fund case (3.28) is always less than the two-asset value function (3.22),

$$v_t(\mathbf{z}_t) \geq v_t^F(\mathbf{z}_t), \text{ for all } \mathbf{z}_t,\ t=0,\ldots,T-1. \tag{3.30}$$

By trading in a mutual fund, as opposed to trading in the constituent assets, the investor is losing the flexibility of efficiently trading for tax purposes in the separate assets. In fact, when trading in multiple assets, the agent is endowed with a *portfolio of tax-timing options*. This grants the investor the ability to optimally realize capital losses in assets that have performed poorly, while postponing capital gains in assets that have performed well. The holder of a mutual fund, on the other side, is forced to trade the two assets in a fixed proportion, and is therefore unable to take advantage of the distribution of embedded gains and losses. In the following we will interpret this difference as the value of the tax-timing option available to an investor who is able to trade optimally in the two assets individually in the presence of capital gains taxes.

To further investigate the forces that drive this reduction in value from trading in a mutual fund, we numerically solve the portfolio problem with a mutual fund in the presence of capital gains taxes and compare the results with the two-asset case. We assume that the portfolio manager is offering to maintain the risky asset in the Merton proportions, i.e., the portfolio weights of risky assets that will prevail if we were to solve the problem in a world without taxes or other frictions.

[15]Notice that we allow the mutual fund to perform wash sales on the individual securities trade.

Let ω_i^m indicate the Merton weights. These can be derived by solving the no-tax case and deriving the trades in the state $(x_1, x_2) = (0, 0)$, i.e.,

$$\omega_1^m = \frac{\alpha_1^*(0,0)}{\alpha_1^*(0,0) + \alpha_2(0,0)}, \text{ and } \omega_2^m = 1 - \omega_1^m.$$

The fund manager will hence engage in trades to guarantee that the Merton proportion is maintained through time. This is only one of many policies which would generate the turnover within the securities in the mutual funds that is necessary to illustrate our results. However, this investment policy has a useful interpretation which will become evident in the following sections.

In Figure G.5 we compute the certainty equivalent costs incurred by trading in a mutual fund. As in the previous section, these costs are defined as the amount of additional wealth necessary to make the mutual fund investor indifferent to holding the two constituent assets. In each panel we report the certainty equivalent cost of trading in a mutual fund as a function of holding (x_1) and basis-to-price (θ_1) of asset 1 for a fixed holding and basis-to-price ratio in asset 2. Figure G.5 confirms the fact that the two-asset case is always preferred to the mutual fund case. The two-asset investor enjoys additional flexibility in the realization of capital gains. Notice, moreover that the degree to which this flexibility can be exploited is higher when there are substantial capital gains embedded, as in the North-West corner of each panel. This also is consistent with the higher magnitude of the certainty equivalent costs observed in Panels A and C, where the second asset has an embedded gain.

Comparing Figures G.4 and G.5 we notice that the magnitude of the certainty equivalent cost in the former is much higher than in the latter. We conclude from this that the gain, in utility terms, from being able to invest in two assets is much bigger in comparison to the no tax case than in comparison to the mutual fund case. In other words, in the current parameterization, most of the tax-flexibility benefits are already captured by the mutual fund. Only a small additional benefit is attributable to the flexibility of trading in the two assets independently.

Figure G.6 reports the trades α_m, representing the fraction of liquidating wealth invested by the agent in the mutual fund.[16] The quantity α_m is the actual inflow and outflow of resources in the mutual fund. We see that, for the case of an embedded gain in asset 1, the total investment

[16]The trades α_m reported in Figure G.6 is equivalent to the sum of the trades in the two assets performed by the mutual fund to keep the securities in the desired proportion.

in the fund is non-monotonic in the holdings of asset 1.[17] Over the regions in which a two-asset investor would prefer not to trade in asset 1, we note that the mutual fund investor decreases her sales or increases her purchases of fund units. This is done to provide the fund manager with enough resources to prevent him from being forced to sell in order to deliver the promised fixed proportion. In turn, the mutual fund investor avoids being passed taxes on sales of asset 1. We summarize the above analysis in the following result:

Result 4 (Tax Flexibility in Mutual Funds) *The certainty equivalent cost from trading in a fixed-proportion mutual fund over trading in the two constituent assets is largest for those states in which the two assets have substantial embedded gains. The fund investment policy is non-monotonic in asset holdings.*

The above results are particularly interesting if we slightly reinterpret the two cases. We can view the two-asset case as a third type of fund, where the fund manager can choose to hold the assets in the fund in any proportion. If we assume that this fund manager trades with the objective of maximizing fund-holder utility, then the manager will fully internalize the tax effects of his trades on fund holders. In contrast, the fixed-proportion mutual fund manager can be interpreted as an agent who trades without regard for the tax consequences of his trades on fund holders. Then the certainty equivalent cost reported in Figure G.5 represents the cost incurred by trading in a non tax-efficient mutual fund.

3.4.2 Tax Cost of Mutual Fund Turnover

In this subsection we further investigate the relationship between optimal diversification and capital gains taxes by departing from the mutual fund case and analyzing the portfolio problem of an individual who is investing in a fund that does not force the investor to realize capital gains and taxes at the end of every year. This will enable us to isolate and quantify the tax burden generated by turnover within an actively managed mutual fund.

To address this question we assume that there exist a security in the market whose return R_m is defined as a weighted average of the returns on the existing securities, i.e.,

$$\tilde{R}_m = \omega \tilde{R}_1 + (1 - \omega)\tilde{R}_2, \quad \omega \in [0, 1]. \tag{3.31}$$

[17]The total trades when asset 1 exhibits a loss can also be non-monotonic, depending on the tax status of asset 2. This is not observed in the figure simply because of the chosen state variable values.

Notice that such a security can be thought of as an *index fund* which trades to achieve the desired balance $\omega/1-\omega$ between securities and never distributes dividends generated by such trades.[18] The investor will pay taxes upon realization of capital gains embedded in the index fund unit. In other words, the transactions subject to taxation are only those concerning the fund unit and not the trades within risky security.[19] With a slight abuse of terminology we will then refer to this security as an *Index Fund.*

In what follows, we treat R_m in (3.31) as the only asset available to the investor. Its basis θ_m will be defined as the cost of purchased units, evolving as a moving average as specified in Section 3.2. Let x_m be the fraction of wealth invested in such asset. The portfolio problem faced by the investor can therefore be stated as follows:

$$v^I(x_{m,t},\theta_{m,t}) = \max_{\alpha_m} E_t\left[\beta v^I(x_{m,t+1},\theta_{m,t+1})\left(\frac{1}{\tilde{M}(x_{m,t},\theta_{m,t};\alpha_{m,t})}\right)^{1-\gamma}\right] \tag{3.32}$$

where

$$x_{m,t+1} = (x_{m,t}+\alpha_m)\tilde{R}_{m,t+1}\tilde{M}(x_{m,t},\theta_{m,t};\alpha_{m,t}) \tag{3.33}$$

$$\theta_{m,t+1} = \begin{cases} \frac{\theta_{m,t}x_{m,t}+\alpha_{m,t}\left(1+\chi_{\{\alpha_{m,t}<0\}}(\theta_{m,t}-1)\right)}{(x_{m,t}+\alpha_{m,t})}\frac{1}{\tilde{R}_{m,t+1}} & \text{if} \quad \theta_{m,t}<1 \\ \frac{1}{\tilde{R}_{m,t+1}} & \text{if} \quad \theta_{m,t}>1 \end{cases} \tag{3.34}$$

and $\tilde{M}(x_{m,t},\theta_{m,t};\alpha_{m,t}) = \frac{W(t)}{W(t+1)}$ is similarly determined as in (3.19) and (3.20). [20]

As in the mutual fund case we assume that the index is following the *Merton policy* in determining the balance between the two assets. This implies that the weight ω can be derived starting from the Merton weights obtained in the no-tax case, i.e.,

$$\omega = \frac{\alpha_1^*(0,0)}{\alpha_1^*(0,0)+\alpha_2^*(0,0)}.$$

[18]As already mentioned, this is an analogy for illustrative purposes only. Real-world tax treatment of index funds is much more complex than our model allows. However, index funds are mutual funds with a low portfolio turnover. Managers in index funds don't buy and sell stocks very often and therefore minimize the effect of taxation. Our index fund is a very active mutual fund which never passes gains or losses to the investor.

[19]Although our analogy between the security R_m and an index fund is only for pedagogical purposes, we point out that a recent proposal at the US Congress advocates a change in the tax treatment of unrealized capital gains for investors in a mutual fund that parallels the example we discuss in this subsection (See Vice Chairman Jim Saxton's proposal, JEC, June 2000, http://www.house.gov/jec/).

[20]Notice that we allow the index fund to perform wash sales on the individual securities trade.

Given the same starting state, and the same fund investment α_m, this rule will generate the same amount of turnover as in the mutual fund case analyzed previously. The difference is now that such turnover will not have tax consequences for the shareholder. Only purchases and sales α_m of the fund unit will be subject to taxation.

Using the same base case parameters in Section 3.2, we solve the portfolio problem for the index fund investor and compare the solution to the ones obtained previously for (i) the two-taxable asset case and (ii) the mutual fund case. We start our analysis by first comparing the index fund case with the two-asset case. Note that this can also be reinterpreted as a comparison between a a tax-efficient fund (the two-asset case) with a fund that trades in order to guarantee a desired balance between assets and *does not* pass gains or losses to its shareholders. Figure G.7 reports the certainty equivalent cost of trading in the index fund as opposed to trading individually in the two constituent assets. Panel A refers to the case in which asset 1 has an embedded gain while Panel B refers to the case in which asset 1 has an embedded loss. Each of the three lines in each panel represent the certainty equivalent cost of trading in an index fund as opposed to trading in the two separate assets for different initial holdings of the two assets. A positive cost means that the two-asset case is relatively more valuable than the index fund case. The comparison between value functions in these two cases is performed by considering that the investment in the index fund is given by

$$x_m = x_1 + x_2$$

and that the basis to price ratio for the index fund is represented by

$$\theta_m = \frac{\theta_1 x_1 + \theta_2 x_2}{x_1 + x_2}. \tag{3.35}$$

In comparing these value functions, we can examine the effects of two main forces. First, the investor in the index fund is unable to reap tax benefits by timing trades individually in the two assets, just as we found in the mutual-fund case. On the other side, trades inside the fund necessary to maintain a target ratio of the two risky assets can be conducted tax-free. Therefore, the index-fund investor is able to more efficiently maintain diversification between the risky assets. The relative balance between the first *tax-timing* factor and the second *diversification* factor explains the evolution of the certainty equivalent costs depicted in Figure G.7. The fact that certainty equivalent costs can be positive or negative suggests that there is no dominance between the two-asset case and the index fund case, as opposed to what we observed for the mutual fund case. We summarize here the properties of the certainty equivalent cost of trading in an index fund compared to trading in the constituent assets.

For the case where asset 1 has a gain (Panel A) we notice that, when the portfolio is tilted towards asset 1 (i.e., line c), the certainty equivalent cost for investing in an index fund increases relative to the two-asset case as the basis of asset 2 increases. This effect is caused by the interaction of diversification benefits and tax-efficient realization. When the second asset has an embedded gain (low θ_2), the index fund is preferred because it allows the investor to attain a diversified portfolio in the risky assets without incurring the distribution of taxable gains. On the other hand, as asset 2 moves towards a loss situation, the ability to manipulate individual holdings, and therefore optimally realize the loss in asset 2 begins to dominate. The only way the index fund holder can take advantage of an embedded loss is to sell units in the fund. However, since the basis for the index fund as a whole is a combination of the bases of the underlying assets (see (3.35)), in Panel A the basis θ_m for the mutual fund will be driven down by the gain embedded in in asset 1.

For the case in which asset 1 has an embedded loss (Panel B) and the portfolio is tilted towards asset 1 (line c), the certainty equivalent cost for the index portfolio is higher when asset 2 has an embedded gain, and lower when asset 2 has an embedded loss (although it is always positive, indicating that the two-asset case is better than index fund case). This happens because the fund is selling with a "weighted-average" basis that becomes more favorable as asset 2 moves into a loss. The two-asset investor is selling only asset one, which has an embedded constant loss.

For the case where the holding of asset 2 is much higher than the holding of asset 1 (line a), the certainty equivalent cost for index trading increases as as the loss in asset 2 increases, independently of the tax status of asset 1 (Panel A or B). This is because diversification becomes less costly for the two-asset investor as the loss in asset 2 increases, so the ability to realize the loss begins to dominate.

When the holdings of the two assets are identical (line b), the certainty equivalent cost for the index fund trader is always positive. In this case the two-asset case is always superior, because there is no gain to be had from the easier diversification in the index-fund.

Finally, we note that when $\theta_1 > 1$ and $\theta_2 > 1$ (implying $\theta_m > 1$) in Panel B, all the lines converge. Again, this is not surprising given that the presence of wash sales allows the investor to neutralize asset holdings and trade to an optimal position.

We conclude this section by looking at the certainty equivalent costs from trading in an index fund versus trading in a mutual fund. We recall that in our terminology, the index fund refers to a mutual fund which does not distributes dividends to shareholders. Moreover, since we impose that both the index and the mutual fund managers have the same motive for trading, we have that these instruments generate the same amount of turnover, conditional on fund investments. The comparison between mutual fund and index fund therefore will capture the tax effect of mutual fund turnover.

Figures G.8 reports the certainty equivalent cost of trading in an index fund as opposed to trading in a mutual fund. We interpret this cost as the tax cost of mutual fund turnover. As in Figure G.7, Panel A refers to the case of a gain in asset 1 and Panel B refers to the case of a loss in asset 1. A positive number means that the mutual fund is preferred to the index fund. In each panel we report the cost for three different balances in the risky assets (x_1, x_2). In what follows we summarize the finding concerning the tax cost of turnover.

We notice that the index fund is better when the asset holdings are unbalanced, and the overweighted asset has a gain (Panel A, line c and Panel B, line a). This captures the fact that an index fund is providing costless diversification. When the two assets are unbalanced and the overweighted asset has a loss (Panel B, line c) the mutual fund is preferred to the index fund. Here both types of funds can provide diversification but the mutual fund also distributes a tax rebate. Finally, when the two assets are balanced (line b in both panels), the mutual fund is preferred. Diversification does not encourage trades, but tax timing in the two assets is more controllable in the mutual fund.

We summarize the above discussion in the following.

Result 5 (Tax Cost of Mutual Fund Turnover) *The index fund is neither dominates nor is dominated by investing in the two assets individually or investing in a mutual fund. The index funds avoids capital gains taxes while maintaining a diversified portfolio but cannot capture the tax benefit of rebalancing at a loss.*

3.5 Wash Sales Constraints

In the above analysis we have assumed that investors have the ability to immediately realize all capital losses and therefore reset the basis before rebalancing their portfolios. In this section we more closely model the tax code by disallowing these wash sales. The imposition of this constraint may cause the investor to delay the realization of losses when the diversification costs of selling are too large.

The restriction on wash sales affects the way in which the basis evolves over time and, as a consequence, the way in which capital gains and losses are computed. Given that now the investor is constrained to only one trade at a time on any security, there is no need to distinguish between an embedded loss and an embedded gain in the description of the evolution of the basis. Using the notation of section 3.2 we define the realized capital gains and losses from a trade $\Delta_i(a,t)$, $a \in \mathcal{F}_t$, $i = 1, 2$ as follows

$$G_i(a,t) = -\Delta_i(t)\left(1 - \frac{B_i(a,t)}{X_i(a,t)}\right)\chi_{\{\Delta_i(a,t)<0\}}, \quad i = 1, 2. \tag{3.36}$$

Equation (3.36) is the equivalent of (3.5) when wash-sales are not allowed. Similarly, the evolution of the bases is

$$B_i(a,t+1) = B_i(\mathbf{a}(a),t) + \Delta_i(\mathbf{a}(a),t)\left(1 + \chi_{\{\Delta_i(\mathbf{a}(a),t)<0\}}\left(\frac{B_i(\mathbf{a}(a),t)}{X_i(\mathbf{a}(a),t)} - 1\right)\right), \quad i = 1, 2, \tag{3.37}$$

where, we recall, $\mathbf{a}(a) \in \mathcal{F}_t$ denotes the set of successors of $a \in \mathcal{F}_{t+1}$. Again (3.37) is the no-wash sales equivalent of (3.8).

An interesting consequence of disallowing wash-sales is that the set of feasible trades $(\Delta_1, \Delta_2, \Delta_Y)$ is no longer convex. We show this in Appendix D together with the fact that allowing wash sales *convexifies* the feasibility set.

In the rest of this section we numerically solve the model by imposing the wash-sale restriction. We analyze the degree in which disallowing wash sales affects investment strategies and the value of the tax-timing option.

3.5.1 Results

Figure G.9 is the no wash sale equivalent of Figure G.1 in Section 3.3. It reports the trades in asset 1 for various values of the state variables. We use a higher basis-to-price ratio ($\theta_i = 1.5$, $i = 1, 2$)

to emphasize that the basis is no longer reset each period, and will therefore take on a wider range of values.

In comparing this case to the case in which wash sales are allowed, we find three interesting differences. First, we observe that the investment strategy in the presence of a loss ($\theta_1 = 1.5$) is characterized by a discontinuity. As the asset holdings increase, the investor abruptly switches from buying to selling. This is explained by the non-convexity of the budget set and concavity of the utility function.

Secondly, the investor with a loss both buys more and sells more than the investor who does not face the wash sale constraint.[21] The wash sale-constrained loss investor, when choosing not to sell for diversification reasons, trades to a position with a larger fraction of wealth invested in asset 1 in comparison to the no-wash sale investor. This happens because the holding embeds a capital loss which is a *cash-like* asset. The loss component of the holdings moves inversely to the asset return. This improves the risk-return profile of asset 1 and therefore induces the investor to hold a higher proportion of wealth in that asset. When the wash sale-constrained loss investor chooses to sell, he moves to a position with a lower fraction of wealth invested in asset 1 in comparison to the unconstrained investor. The unconstrained investor is able to sell his entire portfolio, and then rebalance to an optimal holding. His final portfolio rebalancing decision is therefore unaffected by tax motives. The wash sale-constrained investor, on the other side, has two objectives in choosing a rebalancing strategy: realize as much of the tax loss as possible, and achieve an efficiently diversified portfolio. The first motive pushes the final holdings below the level attained by the unconstrained investor.

Finally, notice that in Panel C, the wash sale-constrained loss investor chooses to dump his entire holding of asset 1 whenever he chooses to sell. This is as close as this investor can come to performing a wash sale. If he was not constrained by tax laws, this investor would repurchase asset 1 as demonstrated in Panel C of Figure G.1. Unable to execute this strategy, the investor must sacrifice diversification to take advantage of the tax-rebate. The bound in the amount of sales here is represented by the short-selling constraint. For the asset with a loss we can say that, in the absence of wash sales, the trade "jumps" from positive to negative, with the discontinuity occurring when the holding in the security is high enough to justify selling for tax reasons.

[21]Note that, as discussed earlier, the trading strategy is identical in the wash sale case for every loss $\theta_i \geq 1$. Therefore, the loss lines in Figure G.1 and Figure G.9 are directly comparable.

An intuitive way of summarizing the above results can be the following:

Result 6 (Trades Under Wash Sale Constraint) *The trading strategy is characterized by (i) a "No-trade region" if the asset has an embedded gain, (ii) a "jump" from buying to selling if the asset has an embedded loss. Overall the "volume" of trades generated in the capital loss case is higher than in the capital gain case (i.e., the investor buys more and sells more when the asset has an embedded loss).*

Figure G.10 is the no wash sale equivalent of Figure G.4 Panel D in Section 3.3. It reports the certainty equivalent gain from trading in the two taxable assets as opposed to trading in two non-taxable assets. The most important thing to note from comparing these two figures, is that when wash sales are prohibited, the dominance of the taxable case breaks down. The constrained investor must weigh the diversification and tax motives in both the loss and the gain states. Comparing the taxable and non-taxable investor with the same liquidating wealth, the taxable investor may be unable to entirely realize the loss, and this translates to a lower utility level.

Finally, Figure G.11 is the no wash sale equivalent of Figure G.5. In Panel C, it is interesting to note that the two asset case is better than the mutual fund when the two assets have opposite tax status. This is because the two-asset investor has the flexibility to realize losses on one asset while deferring gains on the other. The mutual fund investor has less flexibility due to the requirement for fixed proportions. This effect does not show up for the investor who can perform wash sales, because this investor gets an extra opportunity to realize losses both in the fund and two-asset case.

3.6 Conclusions

We solve the portfolio allocation problem with capital gains taxes for an investor with constant relative risk aversion facing an investment set composed of two risky assets and a riskless asset. We interpret the effective tax rate as an endogenous form of transaction cost on sales and thereby formulate the model as a traditional portfolio problem with state dependent transaction costs. The optimal trading strategies are derived by solving a Dynamic Programming problem in discrete time and finite horizon.

In particular, by looking at the optimal trading strategies we characterize a regions of "no trade" similar to what we find in the traditional transaction costs literature. The shape of that region is, however, hard to characterize analytically and we rely on numerical analysis to grasp some intuition. We show that the effect of capital gains taxes on optimal diversification is quite dramatic in the two asset case and that with short-sale restrictions it is sometimes optimal to realize capital gains. We solve our model both in the case where wash sales are admitted, and when wash sales are prohibited by law.

We provide estimate of the value of the tax deferral option by computing the certainty-equivalent cost of incurring taxation on accrual rather than taxation on realization. We show that the value of the tax-timing option increases with embedded gains, and decreases with volatility, and correlation between the assets.

To further examine the impact of the tax-timing option, we compare the two-asset case with two stylized forms of asset funds: an open ended mutual fund which passes gains to investors, and a tax-efficient index fund which is able to rebalance without incurring taxes. The first instrument allows us to characterize the value of the flexibility option available to the two fund investor. The second fund is used to provide a measure of the cost of mutual fund turnover.

Our model is a flexible framework within which we can examine many aspects of the investment problem under taxation. In future research, we will examine the dichotomy between growth and income stocks by altering the level of dividends payed on the two assets. We will also more accurately model the tax code by imposing offsetting requirements on the realization of losses, asymmetry in the treatment of capital gains and losses, and differentiation between long-term and short-term capital gains.

Bibliography

[1] Akian, M., J. L. Menaldi and A. Sulem (1996). "On an Investment-Consumption Model with Transaction Costs." *SIAM Journal of Control and Optimization*, 31, 1, 329-364.

[2] Barclay, M. J., N. D. Pearson and M. S. Weisbach. (1998). "Open-End Mutual Funds and Capital-Gains Taxes." *Journal of Financial Economics*, 49, 3-43.

[3] Bellman, R. (1957). *Dynamic Programming.* Princeton, Princeton Univesity Press.

[4] Berk, J., R. C. Green and V. Naik (1999). "Valuation and Return Dynamics of New Ventures." *NBER Working Paper # 6745.*

[5] Black, F. and M. Scholes (1973). "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81, 637-659.

[6] Brennan, M. and E. Schwartz (1985). "Evaluating Natural Resources Investments." *Journal of Business* 58, 135-157.

[7] Cass, D. and J. Stiglitz (1970). "The Structure of Investor Preference and Asset Returns, and Separability in Portfolio Allocation: A Contribution to the Pure Theory of Mutual Funds." *Journal of Economic Theory*, 2, 122-160.

[8] Childs, P.D. and A.J. Triantis (1997). "Dynamic R&D Investment Policies." *Working Paper*, College of Business and Management and University of Maryland.

[9] Costantinides, G. (1983). "Capital Market Equilibrium with Personal Tax." *Econometrica*, 51, 611-636.

[10] Costantinides, G. (1984). "Optimal Stock Trading with Personal Taxes." *Journal of Financial Economics*, 13, 65-89.

[11] Dammon, R. M., K. Dunn and C. S. Spatt (1989). "A Reexamination of the Value of Tax Options." *Review of Financial Studies*, 2, 341-372.

[12] Dammon, R. M., and C. S. Spatt (1996). "The Optimal Trading and Pricing of Securities with Asymmetric Capital Gains Taxes and Transaction Costs." *Review of Financial Studies*, 9, 921-952.

[13] Dammon, R. M., C. S. Spatt and H. H. Zhang (2000). "Optimal Consumption and Investment with Capital Gains Taxes." *Review of Financial Studies*, forthcoming.

[14] Davis, M. H. A. and A.R. Norman (1990). "Portfolio Selection with Transaction Costs." *Mathematics of Operations Research*, Vol. 15, No. 4, 676-713.

[15] Dixit, A. (1991). "A Simplified Treatment of Some Results Concerning Regulated Brownian Motion." *Journal of Economic Dynamics and Control*, 15, 657-674.

[16] Dixit A. and R. Pindyck (1994). *Investment Under Uncertainty.* Princeton, Princeton University Press.

[17] DSTI (2000) "Differences in Economic Growth across the OECD in the 1990s: The Role of Innovation and Information Technologies. DSTI Preliminary Findings and Contribution to the Growth Project" Paris, OECD, 29 February - 2 March 2000.

[18] Dumas, B. (1991). "Super-Contact and Related Optimality Conditions." *Journal of Economic Dynamics and Control*, 15, 675-686.

[19] Dybvig, P. H., H. K. Koo (1996). "Investment with Taxes." *Washington University in St. Louis Working Paper.*

[20] Fisher, I. (1906). *The Nature of Capital and Income.*

[21] Fisher, I. (1930). *The Theory of Interest.* New York, MacMillan.

[22] Filar, J. and K. Vrieze (1997). *Competitive Markov Decision Processes.* Springer-Verlag, New York.

[23] Fudenberg, D., R. Gilbert, J. Stiglitz and J. Tirole (1983). "Preemption, Leapfrogging and Competition in Patent Races." *European Economic Review* 22, 3-31.

[24] Grenadier, S. (1996a). "The Strategic Exercise of Options: Development Cascades and Overbuilding in Real Estate Markets." *Journal of Finance* 51, 1653-1679.

[25] Grenadier, S. (1999a). "Information Revelation Through Option Exercise." *Review of Financial Studies* 12, 95-130.

[26] Grenadier, S. (1999b). "Option Exercise Games: An Application to the Equilibrium Investment Strategies of Firms." *Working Paper*, Graduate School of Business, Stanford University.

[27] Grenadier, S. and A.M. Weiss (1997). "Investment in technological innovations: An option pricing approach." *Journal of Financial Economics* 44, 397-416.

[28] Grossman, G.M. and C. Shapiro (1987). "Dynamic R&D Competition." *The Economic Journal* 97, 372-387.

[29] Hakansson, N.H. (1970). "Optimal Investment and Consumption Strategies Under Risk for a Class of Utility Functions." *Econometrica* 38, 587-607.

[30] Harris, C. and J. Vickers (1985). "Perfect Equilibrium in a Model of a Race." *Review of Economic Studies* 52, 193-209.

[31] Harris, C. and J. Vickers (1987). "Racing with Uncertainty." *Review of Economic Studies* 54, 1-21.

[32] He, H. (1990). "Convergence from Discrete- to Continuous-Time Contingent Claims Prices." *Review of Financial Studies*, 3, 4, 523-546.

[33] Judd, K. (1985). "Closed-loop Equilibrium in a Model of a Race." *Discussion Paper No. 647.* Management Economics and Decision Science, Kellogg Graduate School of Management, Northwestern University.

[34] Lambrecht, B.M. (1997). "Strategic Sequential Investments and Sleeping Patents." *University of Cambridge, JIMS Working Paper.*

[35] Lambrecht, B.M. and W.R.M. Perraudin (1994). "Option Games." *University of Cambridge, DAE Working Paper No. 9414.*

[36] Lambrecht, B.M. and W.R.M. Perraudin (1997). "Real Options and Preemption." *University of Cambridge mimeo*, Cambridge.

[37] Lee, T. and L. Wilde (1980). "Market Structure and Innovation: a Reformulation." *Quarterly Journal of Economics* 94, 429-436.

[38] Leland, H. (2000). "Optimal Portfolio Implementation with Transaction Costs and Capital Gains Taxes." *University of California at Berkeley. Haas School of Business Working Paper.*

[39] Loury, G. (1979). "Market Structure and Innovation." *Quarterly Journal of Economics* 93, 395-410.

[40] Magill, M., and G. Constantinides (1976). "Portfolio Selection with Transaction Costs." *Journal of Economic Theory*, 13, 245-263.

[41] McDonald, R. and D. Siegel (1986). "The Value of Waiting to Invest." *Quarterly Journal of Economics* 101, 707-727.

[42] Merton, R. C. (1971). "Optimum Consumption and Portfolio Rules in a Continuous-Time Model." *Journal of Economic Theory*, 2, 373-413.

[43] Merton, R. (1973). "The Theory of Rational Option Pricing." *Bell Journal of Economics and Management Science* 4, 141-183.

[44] Myers, S.C. (1977). "Determinants of Corporate Borrowing." *Journal of Financial Economics* 5, 147-176.

[45] OECD (1998) *Technology, Productivity, and Job Creation*, Organisation for Economic Cooperation and Development (OECD).

[46] OECD (1999) *Science, Technology and Industry Scoreboard 1999*, Paris.

[47] Osborne, M.J. and A. Rubinstein (1994). *A Course in Game Theory.* MIT Press.

[48] Puterman, M.L. (1994). *Markov Decision Processes.* John Wiley & Sons, Inc.

[49] Reinganum, J. (1981). "Dynamic Games of Innovations." *Journal of Economic Theory* 25, 21-41.

[50] Reinganum, J. (1982). "A Dynamic Game of R and D: Patent Protection and Competitive Behavior" *Econometrica* 50, 671-688.

[51] Schwartz, E. and M. Moon (1995). "Evaluating Research and Development Investments." *Working Paper* UCLA.

[52] Shreve S. E. and H.M. Soner (1994). "Optimal Investment and Consumption with Transaction Costs." *The Annals of Applied Probability*, Vol. 4, No. 3, 609-692.

[53] Smit, H.T.J. and L.A. Ankum (1993). "A Real Options and Game-Theoretic Approach to Corporate Investment Strategy Under Competition." *Financial Management* 22, 241-250.

[54] Smit, H.T.J. and L. Trigeorgis (1993). "Flexibility and Commitment in Strategic Investment." *Working Paper.* Tinbergen Institute, Erasmus University, Rotterdam.

[55] Smets, F. (1991). "Exporting versus FDI: The Effect of Uncertainty, Irreversibilities and Strategic Interactions." *Working Paper.* Yale University.

[56] Taksar, M., M. Klass, and D. Assaf (1988). "A diffusion Model for Optimal Portfolio Selection in the Presence of Brokerage Fees." *Mathematics of Operations Research*, 13, 277-294.

[57] Trigeorgis, L. (1990). "Valuing the Impact of Uncertain Competitive Arrivals on Deferrable Real Investment Opportunities." *Working Paper.* Boston University.

[58] Trigeorgis, L. (1996). *Real Options.* Cambridge (MA): MIT Press.

[59] von Neumann, J. and O. Morgenstern (1947). *Theory of Games and Economic Behavior.* Princeton, Princeton Univesity Press.

[60] Wang, T. (1999). "Equilibrium with New Investment Opportunities." *Journal of Economic Dynamics and Control*, Forthcoming.

Appendix A

Proofs of Propositions in Chapter 2

Proof of Proposition 2.2.1

Let $s = 1$. From (3.4) and (2.3)

$$E_t\left[\frac{z(t+1)}{z(t)}\delta(t+1)\right] = \delta(t)E_t\left[\exp\left(\mu - r - \frac{1}{2}(\sigma_\delta^2 + \sigma_z^2) + \sigma_\delta\varepsilon_\delta - \sigma_z\varepsilon_z\right)\right].$$

Using the properties of the log-normal distribution we obtain

$$E_t\left[\frac{z(t+1)}{z(t)}\delta(t+1)\right] = \delta(t)\exp(\mu - r - \rho\sigma_\delta\sigma_z) = \delta(t)\beta,$$

where $\rho = cov(\varepsilon_\delta, \varepsilon_z)$, $\beta = \exp(\mu - r - \lambda)$ and $\lambda = \rho\sigma_\delta\sigma_z$.

The proof for $s > 1$ follows trivially. ■

Proof of Proposition 2.3.1

If firm A completes first, then

$$V^A(N, m, \delta(t)) = \sum_{s=0}^{\infty} E_t\left[\frac{z(t+s)}{z(t)}\delta(t+s)\right] = \frac{\delta(t)}{1-\beta}$$

where the last equality follows from Proposition 2.2.1 using the formula for an infinite geometric series. ■

Proof of Proposition 2.3.2

Let $s \equiv (n, m, \delta) \in \mathbb{N}^2 \times \mathbb{R}_+$ indicate a state of the game. By Theorem 4.6.4 in Filar and Vrieze (1997), every nonzero-sum stochastic game has an equilibrium (u, ν) in mixed strategies.[1]

[1]The proof of this result is an application of the Kakutani's fixed point theorem.

Moreover, by Theorem 4.6.5 in Filar and Vrieze (1997), a pair of strategies (u_*, ν_*) is an equilibrium point or the R&D game if, and only if, *for every state* s the pair $(u_*(s), \nu_*(s))$ is a Nash Equilibrium in the subgame originating in state s with payoffs given by

$$\Gamma^A(s, u(s), \nu(s)) = \begin{cases} E^{Q,\Pi}[V^A_{(u,\nu)}(\tilde{s})|s, \nu(s)] & \text{if } u(s) = 0 \\ E^{Q,\Pi}[V^A_{(u,\nu)}(\tilde{s})|s, \nu(s)] - a - b\delta & \text{if } u(s) = 1 \end{cases}$$

$$\Gamma^B(s, u(s), \nu(s)) = \begin{cases} E^{Q,\Pi}_s[V^B_{(u,\nu)}(\tilde{s})|s, u(s)] & \text{if } \nu(s) = 0 \\ E^{Q,\Pi}[V^B_{(u,\nu)}(\tilde{s})|s, u(s)] - a - b\delta & \text{if } \nu(s) = 1 \end{cases}$$

where $\tilde{s}$ denotes the future state of the game whose probability depends on Q, is the equivalent martingale measure and Π, the transition probability matrix described in Table I.[2] For every $s = (n, m, \delta)$ the equilibrium payoffs in pure strategies under the optimal strategy (u_*, ν_*) is therefore characterized by the following system of recursions

$$\begin{aligned} V^A_{(u_*,\nu_*)}(s) = \quad & \max\Big\{\pi^A E^Q_\delta[V^A_{(u_*,\nu_*)}(n+1, m, \tilde{\delta})] + \nu_*(s)\pi^B E^Q_\delta[V^A_{(u_*,\nu_*)}(n, m+1, \tilde{\delta})] \\ & +(1 - \pi^A - \nu_*(s)\pi^B)E^Q_\delta[V^A_{(u_*,\nu_*)}(n, m, \tilde{\delta})] - a - b\delta, \\ & \nu_*(s)\pi^B E^Q_\delta[V^A_{(u_*,\nu_*)}(n, m+1, \tilde{\delta})] + (1 - \nu_*(s)\pi^B)E^Q_\delta[V^A_{(u_*,\nu_*)}(n, m, \tilde{\delta})]\Big\} \end{aligned} \quad (A.1)$$

$$\begin{aligned} V^B_{(u_*,\nu_*)}(s) = \quad & \max\Big\{\pi^B E^Q_\delta[V^B_{(u_*,\nu_*)}(n, m+1, \tilde{\delta})] + u_*(s)\pi^A E^Q_\delta[V^B_{(u_*,\nu_*)}(n+1, m, \tilde{\delta})] \\ & +(1 - u_*(s)\pi^A - \pi^B)E^Q_\delta[V^B_{(u_*,\nu_*)}(n, m, \tilde{\delta})] - a - b\delta, \\ & u_*(s)\pi^A E^Q_\delta[V^B_{(u_*,\nu_*)}(n+1, m, \tilde{\delta})] + (1 - u_*(s)\pi^A)E^Q_\delta[V^B_{(u_*,\nu_*)}(n, m, \tilde{\delta})]\Big\} \end{aligned} \quad (A.2)$$

The pure strategy equilibrium in state s is derived by inspection from (A.1) and (A.2) as follows

$$u_*(s) = \begin{cases} 1 & \text{if } \pi^A E^Q_\delta[V^A_{(u_*,\nu_*)}(n+1, m, \tilde{\delta}) - V^A_{(u_*,\nu_*)}(n, m, \tilde{\delta})] > a + b\delta \\ 0 & \text{if } \pi^A E^Q_\delta[V^A_{(u_*,\nu_*)}(n+1, m, \tilde{\delta}) - V^A_{(u_*,\nu_*)}(n, m, \tilde{\delta})] < a + b\delta \end{cases}, \quad (A.3)$$

and

$$\nu_*(s) = \begin{cases} 1 & \text{if } \pi^B E^Q_\delta[V^B_{(u_*,\nu_*)}(n, m+1, \tilde{\delta}) - V^B_{(u_*,\nu_*)}(n, m, \tilde{\delta})] > a + b\delta \\ 0 & \text{if } \pi^B E^Q_\delta[V^B_{(u_*,\nu_*)}(n, m+1, \tilde{\delta}) - V^B_{(u_*,\nu_*)}(n, m, \tilde{\delta})] < a + b\delta \end{cases}. \quad (A.4)$$

[2]Recall that $u = 1$ means 'invest' for firm A and $u = 0$ means 'wait'. Similarly, $\nu = 1$ means 'invest' for firm B and $\nu = 0$ means 'wait'.

Since δ takes values in a continuum state space ($\mathbb{R}_+$), from the above inequalities, we deduce that the set of mixed strategies is a $P-null$ set. We can see this from the point of view of firm A. Firm A randomizes at s if

$$\pi^A E_\delta^Q[V^A_{(u_*,\nu_*)}(n+1,m,\tilde{\delta}) - V^A_{(u_*,\nu_*)}(n,m,\tilde{\delta})] = a + b\delta$$

In Proposition 2.3.6, we show that the left hand side is a non-decreasing, continuous function of δ. Therefore, since δ takes value in a continuum state space, the above equality is *almost surely* never true. This shows that the set of mixed strategies is *P-null.* Since this holds for every state (n, m, δ), we can rewrite the equilibrium payoff recursions (A.1)-(A.2) as in (2.11)-(2.12) and the optimal decision rules (A.3)-(A.4) as in (2.17)-(2.18). Note that, from (2.17) and (2.18), for every $s = (n, m, \delta)$, only one of the possible combination of actions $(u_*(s), \nu_*(s))$ can be a Nash equilibrium. This shows that, for every state s, the equilibrium in pure strategy is unique.

To obtain the boundary conditions (2.13)-(2.14), we need to compute the value of a completed project. From the expressions of δ in (3.4) and of the pricing kernel in (2.3) and by the properties of log-normal distributions, it is straightforward to show that, for all t,

$$E_t\left[\sum_{s=0}^{\infty}\frac{z(t+s)}{z(t)}\delta(t+s)\middle|\,\delta(t)\right] = \frac{\delta(t)}{1-\beta}.$$

Boundary conditions (2.15)-(2.16) are a consequence of the winner-take-all provision in the R&D game. ■

Proof of Proposition 2.3.5

In symmetric races, $V^B(m,n) = V^A(n,m) \equiv V(n,m)$. If $n = m$, it is immediate to see from (2.17) and (2.18), that neither $[I, W]$ nor $[W, I]$ can be true for any value of δ. ■

Proof of Proposition 2.3.6

Given the strategy of the opponent, the value of the R&D project is analogous to the price of a *compound option* written on the underlying cash flow δ and with exercise price equal to the investment cost. A firm which has completed $n + 1$ stages holds a compound option with a lower exercise price (i.e. lower expected cost to completion) than a firm that has completed n stages. Being such options of the call-type, (i) follows directly. To prove (ii) we use symmetry by rewriting

$V(n,m,\delta) \equiv V^A(n,m) = V^B(m,n)$. Applying (i) to firm B, we obtain (ii). (iii) follows trivially since, by (2.13), $V(n,m,\delta) \leq \frac{\delta}{1-\beta}$ for all δ. ■

Proof of Proposition 2.3.8

By proposition 2.3.6, $\Xi_{n,m}(\tilde{\delta}) \geq 0$ and therefore $E^Q[\Xi_{(n,m)}(\tilde{\delta})|\delta]$ is non-negative. By first order stochastic dominance, the conditional distribution (under the equivalent martingale measure) of $\tilde{\delta}$ given δ is increasing in δ. Therefore $E^Q[\Xi_{(n,m)}(\tilde{\delta})|\delta]$ is non-negative and non-decreasing in δ. Since the investment cost, $a + b\delta$, is increasing in δ (at rate b), the result follows immediately from conditions (2.17) and (2.18) describing the optimal investment decision rules. ■

Proof of Proposition 2.3.9

The proof follows immediately from first order stochastic dominance and (2.17) and (2.18). ■

Proof of Proposition 2.3.10

(i) Suppose we solved (backwards) for the value $V(j,k,\delta)$ of the R&D project for all subgames (j,k), $j = N-1,\ldots,n+1$, $k = N-1,\ldots,m+1$, $(n+1,m)$ and $(n,m+1)$. In subgame (n,m), let us keep B's strategy $\nu_*(n,m)$, constant for all δ. If $\nu_* = 0$ for all δ, firm A's value, $V_1(n,m,\delta)$ is obtained by solving the following recursion:

$$\begin{aligned} V_1(n,m,\delta) &= max\Big\{\pi E_\delta^Q[V(n+1,m,\tilde{\delta})] + (1-\pi)E_\delta^Q[V_1(n,m,\tilde{\delta})] - a - b\delta, \\ &\quad E_\delta^Q[V_1(n,m,\tilde{\delta})]\Big\} \\ &\equiv \mathcal{A}^{(n,m)}(V_1(n,m,\delta)) \end{aligned} \tag{A.5}$$

Similarly, if $\nu_* = 1$ for all δ, firm A's value, $V_2(n,m,\delta)$ is obtained by solving the following recursion:

$$\begin{aligned} V_2(n,m,\delta) &= max\Big\{\pi E_\delta^Q[V(n+1,m,\tilde{\delta})] + \pi E_\delta^Q[V(n,m+1,\tilde{\delta})] + \\ &\quad +(1-2\pi)E_\delta^Q[V_2(n,m,\tilde{\delta})] - a - b\delta, \\ &\quad \pi E_\delta^Q[V_2(n,m+1,\tilde{\delta})] + (1-\pi)E_\delta^Q[V_2(n,m,\tilde{\delta})]\Big\} \\ &\equiv \mathcal{B}^{(n,m)}(V_2(n,m,\delta)) \end{aligned} \tag{A.6}$$

Starting from subgame $(N-1, N-1)$ and proceeding backwards, it is easy to note that Blackwell's sufficient conditions (See Stokey and Lucas (1989)) are satisfied for every n, m. This implies that $\mathcal{A}^{(n,m)}$ and $\mathcal{B}^{(n,m)}$ are contraction mappings for all n and m and, therefore, the functional equations (A.5) and (A.6) admit unique fixed points, $V_1(n, m, \delta)$ and $V_2(n, m, \delta)$ respectively, both non-decreasing in δ.[3] By proposition 2.3.8, firm B is adopting a threshold strategy, therefore there exists a $\delta_*(m, n) \leq \infty$ such that $\nu_*(n, m) = 0$ for $\delta < \delta_*$ and $\nu_*(n, m) = 1$ for $\delta \geq \delta_*$. Hence, the resulting value in subgame (n, m) is given by the mixture

$$V(n, m, \delta) = V_1(n, m, \delta)\chi_{\{\delta<\delta_*(m,n)\}} + V_2(n, m, \delta)\chi_{\{\delta\geq\delta_*(m,n)\}}$$

From (A.5) and (A.6) it is clear that

$$V_2(n, m, \delta) = \mathcal{B}^{(n,m)}(V_2(n, m, \delta)) = \mathcal{A}^{(n,m)}(V_2(n, m, \delta)) - \pi E_\delta^Q[V_2(n, m, \delta)] \qquad \text{(A.7)}$$

Since $V_1(n, m, \delta)$ is the unique fixed point of $\mathcal{A}^{(n,m)}$, $\mathcal{A}^{(n,m)}(V_2(n, m, \delta)) \leq \mathcal{A}^{(n,m)}(V_1(n, m, \delta)) = V_1(n, m, \delta)$. Hence, $V_2(n, m, \delta) \leq V_1(n, m, \delta)$ for all δ and therefore $\lim_{\delta\to\delta_*^-} V_1(n, m, \delta) \geq \lim_{\delta\to\delta_*^+} V_2(n, m, \delta)$.

(ii) Let $\mathcal{D}$ be the set of functions (A.13) defined in the proof of Proposition 2.4.1. Since $\mathcal{A}^{(n,m)}$ and $\mathcal{B}^{(n,m)}$ are contractions, for every $f, g \in \mathcal{D}$,

$$V_1(n, m, \delta) = \lim_{k\to\infty} (\mathcal{A}_k^{(n,m)} f)(\delta)$$

and

$$V_2(n, m, \delta) = \lim_{k\to\infty} (\mathcal{A}_k^{(n,m)} g)(\delta),$$

where $(\mathcal{A}_k^{(n,m)} f) = \mathcal{A}^{(n,m)}(\mathcal{A}_{k-1}^{(n,m)} f)$, and $(\mathcal{B}_k^{(n,m)} g) = \mathcal{B}^{(n,m)}(\mathcal{B}_{k-1}^{(n,m)} g)$, $k \geq 2$. In order to characterize the relationship between the fixed point of $\mathcal{A}$ and $\mathcal{B}$, let us look at the sequences $\{f_k\}$ and $\{g_k\}$ where $f_k = \mathcal{A}(f_{k-1})$ and $g_k = \mathcal{B}(f_{k-1})$, $k \geq 1$. Being $\mathcal{A}^{(n,m)}$ and $\mathcal{B}^{(n,m)}$ contraction mappings, we can, without loss of generality, select a common starting point $f_0 = g_0$ for the two sequences. From (A.7), we know that $\mathcal{B}^{(n,m)}(g) = \mathcal{A}^{(n,m)}(g) - \pi E_\delta^Q[g]$ for all $g \in \mathcal{D}$, therefore we obtain

$$\begin{aligned} g_1 &= \mathcal{B}^{(n,m)}(g_0) = \mathcal{A}^{(n,m)}(g_0) - \pi E_\delta^Q[g_0] \\ &= \mathcal{A}^{(n,m)}(f_0) - \pi E_\delta^Q[g_0] \\ &= f_1 - \pi E_\delta^Q[g_0] \end{aligned}$$

[3]See argument in the proof of Proposition 2.4.2 for a complete characterization of monotonicity in δ.

With a little algebra we can show straightforwardly that the two sequences $\{f_k\}$ and $\{g_k\}$ satisfy the following relationship

$$g_k = f_k - \pi \sum_{j=1}^{k} E_{\delta}^{Q}[g_j]. \tag{A.8}$$

Since $V_1(n,m,\delta) = \lim_{k\to\infty} f_k$ and $V_2(n,m,\delta) = \lim_{k\to\infty} g_k$, taking limits on both sides of (A.8) we obtain

$$V_2(n,m,\delta) = V_1(n,m,\delta) - \pi \sum_{j=1}^{\infty} E^{Q}[g_j|\delta]. \tag{A.9}$$

Let $J(\delta) \equiv \pi \sum_{j=1}^{\infty} E^{Q}[g_j|\delta]$. Since $g_j \in \mathcal{D}$ for all j, $J(\delta) \leq \pi \sum_{j=1}^{\infty} \frac{\beta^j \delta}{(1-\beta)} = \pi \frac{\beta\delta}{(1-\beta)^2}$ At δ_*, $J(\delta_*) \leq \pi \frac{\beta\delta_*}{(1-\beta)^2}$ which represents the upper bound of the jump in the value function $V(n,m,\delta)$. ■

Proposition A.0.1 *If there are no fixed costs, i.e., $a = 0$, then the value function solving (2.11)-(2.12) is homogeneous of degree one in δ and we can write*

$$V(n,m,\delta(t)) = h(n,m)\delta(t), \;\; \textit{for all}\; t,n,m. \tag{A.10}$$

where $h(n,m)$ satisfies the following recursion

$$h(n,m) = \begin{cases} 0 & \textit{if} \begin{cases} \pi\beta h(n+1,m) < b \\ \pi\beta h(m+1,n) < b \end{cases} \\[2ex] \pi\beta\theta_1 h(n,m+1) & \textit{if} \begin{cases} \pi\beta(h(n+1,m) - \pi\beta\theta_1 h(n,m+1)) < b \\ \pi\beta h(m+1,n) > b \end{cases} \\[2ex] \pi\beta\theta_1 h(n+1,m) - b\theta_1 & \textit{if} \begin{cases} \pi\beta h(n+1,m) > b \\ \pi\beta(h(m+1,n) - \pi\beta\theta_1 h(m,n+1)) < b \end{cases} \\[2ex] \pi\beta\theta_2(h(n+1,m) + h(n,m+1)) - b\theta_2 & \textit{if} \begin{cases} \pi\beta(h(n+1,m) - \pi\beta\theta_1 h(n,m+1)) > b \\ \pi\beta(h(m+1,n) - \pi\beta\theta_1 h(m,n+1)) > b \end{cases} \end{cases}$$

with

$$\theta_1 = \frac{1}{1-\beta(1-\pi)}, \quad \theta_2 = \frac{1}{1-\beta(1-2\pi)}$$

and boundary conditions $h(N,m) = 1/(1-\beta)$, for all $m < N$, $h(n,N) = 0$, for all $n < N$.

Proof: We solve for the value backwards from $n = N - 1$ and $m = N - 1$. After substituting the conjectured expressions (A.10) of $V(n, m, \delta)$ and $V(m, n, \delta)$ in (2.11) and (2.12) and simplifying, we obtain

$$\begin{aligned} h(n,m) &= (1-\nu_*)\max\{\pi\beta h(n+1,m) + (1-\pi)\beta h(n,m) - b, \beta h(n,m)\} + \\ &+\nu_*\max\{\pi\beta h(n+1,m) + \pi\beta h(n,m+1) + (1-2\pi)\beta h(n,m) - b, \\ &\pi\beta h(n,m+1) + (1-\pi)h(n,m)\} \end{aligned} \tag{A.11}$$

$$\begin{aligned} h(m,n) &= (1-u_*)\max\{\pi\beta h(m+1,n) + (1-\pi)\beta h(m,n) - b, \beta h(m,n)\} + \\ &+u_*\max\{\pi\beta h(m+1,n) + \pi\beta h(m,n+1) + (1-2\pi)\beta h(m,n) - b, \\ &\pi\beta h(m,n+1) + (1-\pi)h(m,n)\} \end{aligned} \tag{A.12}$$

where u_* and ν_* are the strategy of firm A and B respectively. Suppose $\nu_* = 0$ in (A.11). From (2.18), this happens when $h(m+1,n) - h(m,n) < b$ (i.e. B is mothballing). Solving for $h(n,m)$ in (A.11) we obtain

$$h(n,m) = \begin{cases} 0 & \text{if} \quad \pi\beta(h(n+1,m) - h(n,m)) < b \\ \pi\beta\theta_1 h(n+1,m) - b\theta_1 & \text{if} \quad otherwise \end{cases}.$$

Similarly, when $h(m+1,n) - h(m,n) > b$, (i.e. B invests) we obtain

$$h(n,m) = \begin{cases} \pi\beta\theta_1 h(n,m+1) & \text{if} \quad \pi\beta(h(n+1,m) - h(n,m)) < b \\ \pi\beta\theta_2(h(n+1,m) + h(n,m+1)) - b\theta_2 & \text{if} \quad otherwise \end{cases}.$$

Repeating the same procedure for $h(m,n)$, replacing the expressions of $h(n,m)$ and $h(m,n)$ thus obtained in the constraints defining the prevailing investment strategies and noticing that $\pi\beta\theta_2/(1-\pi\beta\theta_2) = \pi\beta\theta_1$, we obtain the recursion stated in the proposition. ■

Remark A.0.2 Note that the function $h(n,m)$ in the above proposition is dependent only on the values $h(n+1,m)$, $h(n,m+1)$, $h(m+1,n)$, $h(m,n+1)$ all of which are known when solving for the value in the subgame (n,m).

Remark A.0.3 The homogeneity property shown in Proposition A.0.1 provides both a *feasibility* condition and an *asymptotic* condition that is used extensively in the solution of the model. The feasibility condition follows from observing that $h(0,0)\delta$ represents an upper bound of the value function $V(0,0,\delta)$ at the beginning of the race. Therefore, if $h(0,0) = 0$ then the project is not economical to undertake. The asymptotic condition relies on the fact that for high values of δ the fixed component of the investment cost become relatively negligible and the value function behaves as one emerging from the no-fixed-cost case. We use this to approximate the value function for high values of cash flows. In Appendix B, we provide a detailed description of the algorithm used to solve for the value function in the general case.

Proof of Proposition 2.4.1

The boundary conditions (2.22) define $V^J(n,m,\delta)$ for $n = N$, $m < N$ and for $m = N$, $n < N$. Let

$$\mathcal{D} = \left\{ f : \sup_{\delta} \frac{|f(\delta)|}{\delta} \leq \frac{1}{1-\beta} \right\} \tag{A.13}$$

be the set of all continuous functions $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ uniformly bounded by the value of the completed project. We define the following operator

$$\begin{aligned}
(\mathcal{T}^{(N-1,N-1)}f)(\delta) &= max\{E_{\delta}^{Q}[f(\tilde{\delta}))], \\
&\quad \pi E_{\delta}^{Q}[V^J(N,m,\tilde{\delta})] + (1-\pi)E_{\delta}^{Q}[f(\tilde{\delta})] - a - b\delta, \\
&\quad \pi E_{\delta}^{Q}[V^J(n,N,\tilde{\delta})] + (1-\pi)E_{\delta}^{Q}[f(\tilde{\delta})] - a - b\delta, \\
&\quad \pi E_{\delta}^{Q}[V^J(N,m,\tilde{\delta})] + \pi E_{\delta}^{Q}[V^J(n,N,\tilde{\delta})] + \\
&\quad (1-2\pi)E_{\delta}^{Q}[f(\tilde{\delta})] - 2(a+b\delta)\}
\end{aligned} \tag{A.14}$$

where $\tilde{\delta} = \delta \exp\{\mu - .5\sigma_{\delta}^2 + \sigma_{\delta}\varepsilon_{\delta}(t+1)\}$. It is evident that $\mathcal{T}^{(N-1,N-1)} : \mathcal{D} \rightarrow \mathcal{D}$. Moreover, if $f' \geq f$ then $\mathcal{T}^{(N-1,N-1)}(f') \geq \mathcal{T}^{(N-1,N-1)}(f)$, since each term of the $\max(\cdot)$ in (A.14) is monotone in f. Finally, by the definition of E_{δ}^{Q}, we note that for $\gamma \geq 0$,

$$\mathcal{T}^{(N-1,N-1)}(f+\gamma) \leq \mathcal{T}^{(N-1,N-1)}(f) + e^{-r}\gamma,$$

therefore there exists a $\xi \in (0,1)$ such that $\mathcal{T}^{(N-1,N-1)}(f+\gamma) \leq \mathcal{T}^{(N-1,N-1)}(f) + \xi\gamma$, $\gamma \geq 0$. The last two properties guarantee that $\mathcal{T}^{(N-1,N-1)}$ is a contraction mapping (Blackwell sufficient conditions), and therefore, it admits a unique fixed point $V^J(N-1,N-1,\delta) \in \mathcal{D}$.

The remainder of the proof follows by induction. Suppose we prove that $\mathcal{T}^{(n+1,m)}$ and $\mathcal{T}^{(n,m+1)}$, are contraction mappings defining fixed points $V^J(n+1,m,\delta)$ and $V^J(n,m+1,\delta)$ respectively. Then, by the arguments above, $\mathcal{T}^{(n,m)}$ is a contraction mapping defining the fixed point $V^J(n,m,\delta)$, $n,m = 0,\ldots,N-1$. ■

Proof of Proposition 2.4.2

(i) In order to prove monotonicity in δ, let $x = a/\delta$. With this change of variable, direct substitution in (2.21) allows us to claim that for all $n = 0,\ldots,N-1$, $m = 0,\ldots,N-1$, $V^J(n,m,\delta) = \delta Gs(n,m,x)$ where $Gs(n,m)$ is the unique fixed point of the following operator

$$
\begin{aligned}
(\mathcal{A}^{(n,m)}f)(x) \quad = \quad & \max\{E_x^Q[f(x'))], \\
& \pi E_x^Q[Gs(n+1,m,x')] + (1-\pi)E_x^Q[f(x')] - x - b, \\
& \pi E_x^Q[Gs(n,m+1,x')] + (1-\pi)E_x^Q[f(x')] - x - b, \\
& \pi E_x^Q[Gs(n+1,m,x')] + \pi E_x^Q[Gs(n,m+1,x')] + \\
& (1-2\pi)E_x^Q[f(x')] - 2(x+b))\}
\end{aligned}
\tag{A.15}
$$

where $x' = a/\tilde{\delta} = x\exp\{-[\mu - .5\sigma_\delta^2 + \sigma_\delta\varepsilon_\delta(t+1)]\}$. Let us consider the case of $n = N-1$ and $m = N-1$. By definition, $Gs(N,\cdot) = Gs(\cdot,N) = 1/(1-\beta)$ which is weakly decreasing in x. If f is decreasing in x, then $(\mathcal{A}^{(N-1,N-1)}f)(x)$ will be decreasing too. This is true since every component of the max in $\mathcal{A}^{(N-1,N-1)}$ are decreasing in x. Since $Gs(N-1,N-1,x)$ is the *unique* fixed point of the contraction mapping $(\mathcal{A}^{(N-1,N-1)}f)$, then $Gs(N-1,N-1,x) = \lim_{k\to\infty}(\mathcal{A}_k^{(N-1,N-1)}f)(x)$ where $(\mathcal{A}_k^{(N-1,N-1)}f) = \mathcal{A}^{(N-1,N-1)}(\mathcal{A}_{k-1}^{(N-1,N-1)}f)$, $k \geq 2$. Therefore we conclude that $Gs(N-1,N-1,x)$ is decreasing in x. Similarly, by induction, we can prove that $Gs(N-1,m,x)$ for $m = 0,\ldots,N-2$ and $Gs(n,N-1,x)$ for $n = 0,\ldots,N-2$ are decreasing in x. Thus we have shown that, if $Gs(N,\cdot,x)$ is decreasing in x then $Gs(N-1,\cdot,x)$ is also decreasing in x and, similarly, if $Gs(\cdot,N,x)$ is decreasing in x then $Gs(\cdot,N-1,x)$ is also decreasing in x. The remainder of the proof follows by induction. Let $n = N-2$ and $m = N-2$. Using (A.15), we can prove that, if f is nondecreasing in x, then $(\mathcal{A}^{(N-2,N-2)}f)(x)$ is decreasing in x and so on. Therefore, we proved that $Gs(n,m,x)$ is monotonically decreasing in x for $n = 0,\ldots,N$, $m = 0,\ldots,N$. Since $V^J(n,m,\delta) = \delta Gs(n,m,x)$

and $x = a/\delta$, we conclude that $V^J(n, m, x)$ is monotonically increasing in δ for $n = 0, \ldots, N$, $m = 0, \ldots, N$.

The proof of convexity follows the same logic. Since (A.14) is a contraction mapping, it is sufficient to show that, if f is convex, then $\mathcal{T}^{(n,m)}$ is convex too. Let us start with $n = N-1$ and $m = N-1$. Since $V^J(n+1, m, \delta) = V^J(N, N-1, \delta) = \delta/(1-\beta)$ and $V^J(n, m+1, \delta) = V^J(N-1, N, \delta) = \delta/(1-\beta)$, inspection of (A.14) shows that if f is convex in δ, then $(\mathcal{T}^{(N-1,N-1)}f)(\delta)$ is also convex in δ, being the maximum of convex functions a convex function. Since $\mathcal{T}^{(N-1,N-1)}$ is a contraction mapping, it admits a unique fixed point $V^J(N-1, N-1, \delta)$ which is convex in δ. Similarly, by induction, we can prove that $V^J(N-1, m, \delta)$ for $m = 0, \ldots, N-2$ and $V^J(n, N-1, \delta)$ for $n = 0, \ldots, N-2$ are convex in δ. Thus we have shown that, if $V^J(N, \cdot, x)$ is convex in δ, then $V^J(N-1, \cdot, \delta)$ is also convex in δ and, similarly, if $V^J(\cdot, N, \delta)$ is convex in δ then $V^J(\cdot, N-1, \delta)$ is also convex in δ. By induction, we conclude that the fixed point $V^J(n, m, \delta)$ of (A.14) is convex in δ for all $n = 0, 1, \ldots, N$ and $m = 0, 1, \ldots, N$.

(ii). We now prove that $V^J(n, m, \delta)$ is increasing in n and m. The proof, again, is by induction. Let $m = N-1$ and consider the family of functions $V^J(n, N-1, \delta)$ as n varies. We know that, by construction, $V^J(N, N-1, \delta) \geq V(N-1, N-1, \delta)$. Suppose that it is true that $V^J(N, N-1, \delta) \geq V^J(N-1, N-1, \delta) \geq \ldots \geq V^J(n+1, N-1, \delta)$ (induction hypothesis) for all δ. We need to show that $V^J(n+1, N-1, \delta) \geq V^J(n, N-1, \delta)$. Let $f \in \mathcal{D}$ be such that $f(\delta) \leq V^J(n+1, N-1, \delta)$ for all δ. Then, from (A.14) we have

$$
\begin{aligned}
(\mathcal{T}^{(n,N-1)}f)(\delta) \;=\; & max\{E^Q_\delta[f(\tilde{\delta}))], \\
& \pi E^Q_\delta[V^J(n+1, N-1, \tilde{\delta})] + (1-\pi)E^Q_\delta[f(\tilde{\delta})] - a - b\delta, \\
& \pi E^Q_\delta[V^J(n, N, \tilde{\delta})] + (1-\pi)E^Q_\delta[f(\tilde{\delta})] - a - b\delta, \\
& \pi E^Q_\delta[V^J(n+1, N-1, \tilde{\delta})] + \pi E^Q_\delta[V^J(n, N, \tilde{\delta})] + \\
& (1-2\pi)E^Q_\delta[f(\tilde{\delta})] - 2(a + b\delta)\}
\end{aligned}
$$

Now, notice that (i) $E^Q_\delta[f(\tilde{\delta}))] \leq V^J(n+1, N-1, \delta)$ by construction, (ii) $E^Q_\delta[V^J(n+1, N-1, \tilde{\delta})] \leq E^Q_\delta[V^J(n+2, N-1, \tilde{\delta})]$ by the induction hypothesis, (iii)$E^Q_\delta[V^J(n, N, \tilde{\delta})] \leq E^Q_\delta[V^J(n+1, N, \tilde{\delta})]$ by definition of $V^J(\cdot, N, \delta)$. Hence, we conclude that $(\mathcal{T}^{(n,N-1)}f)(\delta) \leq (\mathcal{T}^{(n+1,N-1)}V^J(n+1, m)) = V^J(n+1, N-1, \delta)$. Therefore, if $f(\delta) \leq V^J(n+1, N-1, \delta)$, then $(\mathcal{T}^{(n,N-1)})(f) \leq V^J(n+1, N-1, \delta)$. By the same argument as before, since $\mathcal{T}^{(n,N-1)}$ is a contraction mapping with

$V^J(n, N-1, \delta)$ as unique fixed point, it must be the case that $V^J(n, N-1, \delta) \leq V^J(n+1, N-1, \delta)$ for all $n = 0, \ldots, N-1$. The remainder of the proof follows by induction on m. The proof that $V^J(n, m+1, \delta) \geq V^J(n, m, \delta)$ is similar.

(iii) The last part of the proposition follows trivially since

$$V^J(n, m, \delta) \leq \frac{\delta}{1-\beta} \to 0 \text{ as } \delta \to 0, \ \ \forall n, m. \ \blacksquare$$

Proof of Proposition 2.4.3

Let us consider the case in which it is optimal for the social planner to invest in both units. By inspection of (2.21), we notice that this occurs when the following conditions are satisfied

$$\pi E_\delta^Q[V^J(n+1, m, \tilde{\delta}) - V^J(n, m, \tilde{\delta})] \ > \ a + b\delta \qquad (A.16)$$

$$\pi E_\delta^Q[V^J(n, m+1, \tilde{\delta}) - V^J(n, m, \tilde{\delta})] \ > \ a + b\delta \qquad (A.17)$$

Equation (A.16) refers to the decision to activate unit #1 and equation (A.17) refers to the decision to activate unit #2.

From (2.17)-(2.18) the Nash equilibrium $[I, I]$ occurs when

$$\pi E_\delta^Q[V(n+1, m, \tilde{\delta}) - V(n, m, \tilde{\delta})] \ > \ a + b\delta \qquad (A.18)$$

$$\pi E_\delta^Q[V(m+1, n, \tilde{\delta}) - V(m, n, \tilde{\delta})] \ > \ a + b\delta \qquad (A.19)$$

where (A.18) is the decision rule of player A and (A.19) is the decision rule of player B.

Let us consider (A.16) and (A.18). Note that they can be rewritten as

$$\pi E_\delta^Q[\Xi_{(n,m)}^J(\tilde{\delta})] \ > \ a + b\delta$$

$$\pi E_\delta^Q[\Xi_{(n,m)}^A(\tilde{\delta})] \ > \ a + b\delta$$

By Propositions 2.3.6 and 2.4.2, $\Xi_{(n,m)} \geq 0$ and $\Delta_{(n,m)}^J \geq 0$, therefore $E_\delta^Q[\Xi_{(n,m)}(\tilde{\delta})]$ and $E_\delta^Q[\Xi_{(n,m)}^J(\tilde{\delta})]$ are non-decreasing in δ by first order stochastic dominance. Let $\delta_*^J(n, m) = \inf\{\delta : \pi E_\delta^Q[\Xi_{(n,m)}^J(\tilde{\delta})] \geq a + b\delta\}$ and $\delta_*(n, m) = \inf\{\delta : \pi E_\delta^Q[\Xi_{(n,m)}(\tilde{\delta})] \geq a + b\delta\}$

If $\Xi^J_{(n,m)}(\delta) \geq \Delta^A_{(n,m)}(\delta)$ for all δ then

$$\pi E^Q_\delta[\Xi^J_{(n,m)}(\tilde{\delta})] \leq \pi E^Q_\delta[\Xi^A_{(n,m)}(\tilde{\delta})]$$

Since both terms in the above inequality are non-decreasing in δ, it must be the case that $\delta^J_*(n,m) \geq \delta_*(n,m)$. ■

Proposition A.0.4 *If there are no fixed costs, $(a = 0)$, then the value function solving (2.21) is homogeneous of degree one in δ and we can write*

$$V^J(n,m,\delta(t)) = h^J(n,m)\delta(t), \text{ for all } t,n,m,$$

where $h^J(n,m)$ satisfies the following recursion

$$h^J(n,m) = \begin{cases} 0 & \text{if } \begin{cases} \frac{\pi\beta h^J(n+1,m)}{b} < 1 \\ \frac{\pi\beta h^J(m+1,n)}{b} < 1 \end{cases} \\ \pi\beta\theta_1 h^J(n,m+1) - b\theta_1 & \text{if } \begin{cases} \frac{\pi\beta(h^J(n+1,m)-\pi\beta\theta_1 h^J(n,m+1))}{b(1-\pi\beta\theta_1)} < 1 \\ \frac{\pi\beta h^J(m+1,n)}{b} > 1 \end{cases} \\ \pi\beta\theta_1 h^J(n+1,m) - b\theta_1 & \text{if } \begin{cases} \frac{\pi\beta h^J(n+1,m)}{b} > 1 \\ \frac{\pi\beta(h^J(m+1,n)-\pi\beta\theta_1 h^J(m,n+1))}{b(1-\pi\beta\theta_1)} < 1 \end{cases} \\ \pi\beta\theta_2(h^J(n+1,m) + h^J(n,m+1)) - 2b\theta_2 & \text{if } \begin{cases} \frac{\pi\beta(h^J(n+1,m)-\pi\beta\theta_1 h^J(n,m+1))}{b(1-\pi\beta\theta_1)} > 1 \\ \frac{\pi\beta(h^J(m+1,n)-\pi\beta\theta_1 h^J(m,n+1))}{b(1-\pi\beta\theta_1)} > 1 \end{cases} \end{cases}$$

with

$$\theta_1 = \frac{1}{1-\beta(1-\pi)}, \quad \theta_2 = \frac{1}{1-\beta(1-2\pi)}$$

and boundary conditions $h^J(N,m) = 1/(1-\beta)$, for all $m < N$, $h^J(n,N) = 1/(1-\beta)$, for all $n < N$.

Proof of Proposition 2.5.3

From equation (2.4), the cum-dividend value of the cash flow stream $\delta(t)$ is $\delta(t)/(1-\beta)$. According to (2.24), the expected return on the cash flow stream $\delta(t)$ is

$$E_t[1 + R_{t+1}] = \frac{E^P_t[\delta(t+1)]}{\frac{\delta(t)}{1-\beta} - \delta(t)}$$

$$= \frac{\delta(t)\exp(\mu)}{\delta(t)\beta} = \exp(\lambda + r)$$

where the denominator in the first equality is the ex-dividend value of the completed project, the second equality follows from (3.4), using the property of log-normal distributions, and the third from the definition of β in (2.4). ■

Proof of Proposition 2.5.4

From Proposition A.0.1, we know that in the absence of fixed costs $V(n, m, \delta) = h(n, m)\delta$. By definition of equivalent martingale measures, the ex-dividend value of such venture at time t is $E_t^{\Pi,Q}[V(n(t+1), m(t+1), \delta(t+1))]$. Since technical uncertainty is uncorrelated with the pricing kernel, we have

$$E^{\Pi,Q}[V(n(t+1), m(t+1), \delta(t+1))] = \delta(t)\beta E^{\Pi}[h(n(t+1), m(t+1))] \tag{A.20}$$

The numerator in (2.24) is similarly derived as

$$E^{\Pi,P}[V(n(t+1), m(t+1), \delta(t+1))] = \delta(t)\exp(\mu)E^{\Pi}[h(n(t+1), m(t+1))] \tag{A.21}$$

Taking the ratio of (A.20) and (A.21) and simplifying, we obtain the desired result. ■

Appendix B

Algorithm to solve for $V(n, m, \delta)$ in Chapter 2

Our numerical solution relies on an extended version of the value iteration algorithm used in standard dynamic programming problems (see Puterman (1994) and references therein). The main recursion is represented by the system (2.11)-(2.12) along with the boundary conditions (2.13)-(2.16). The interaction between players implies that we need to solve for the value functions $V(n, m, \delta)$ and $V(m, n, \delta)$ *jointly.*

We start from stage $(N-1, N-1)$ and move backwards to stage $(0, 0)$, thus obtaining a *subgame perfect* equilibrium in the patent race. Since we proved in Proposition 2.3.10 that the resulting value function may not be continuous, the actual convergence algorithm is not performed on $V(n, m, \delta)$, but on the expected value under the equivalent martingale measure Q, $E_t^Q[V(n, m, \tilde{\delta})|\delta]$, which is continuous and non-decreasing in δ. In every subgame, we start by guessing a functional form of $E_t^Q[V(n, m, \tilde{\delta})|\delta]$ (the solution of the future games are used as boundary conditions), and derive the value function in the present subgame. We compute the expectation under the equivalent martingale measure Q and compare it with the guess. We iterate until the new value is not more than 0.0001% different from the previous one. Once a value is found, we move backward to the previous subgame.

In computing expectations, we need to perform numerical integration on a discrete grid of values for δ. For values between these grid points, we use quadratic interpolation and, for values outside the grid, we use the expressions for the linear case shown in Proposition A.0.1. This is justified by the fact that, asymptotically, for high values of δ, the value function behaves like the value function in the absence of fixed costs.

The numerical solution for the RJV problem follows the same logic.

Appendix C

Tables and Figures for Chapter 2

Table III. Parameter Values for the Numerical Solution. The table lists the values chosen for all parameters required to solve the the patent race model and the RJV model (quarterly values): the number of phases to complete a project (N), the fixed component of the development cost (a), the variable component of the development cost (b), the probability of successfully overcoming an hurdle (π), the average growth (μ) of the potential cash flow process, the volatility of cash flows (σ_δ), the volatility of innovation in the pricing kernel (σ_z), the correlation between innovations in the cash flow process and in the pricing kernel (ρ) and the riskless interest rate (r). The last entry (λ) is the risk premium earned by the completed project and is derived in Proposition 2.3.1.

Parameter	Value (quarterly)
N	20
a	0.5
b	0.05
π	.25
μ	.0075
σ_δ	.30
σ_z	.20
ρ	0.5
r	.0175
λ	.03

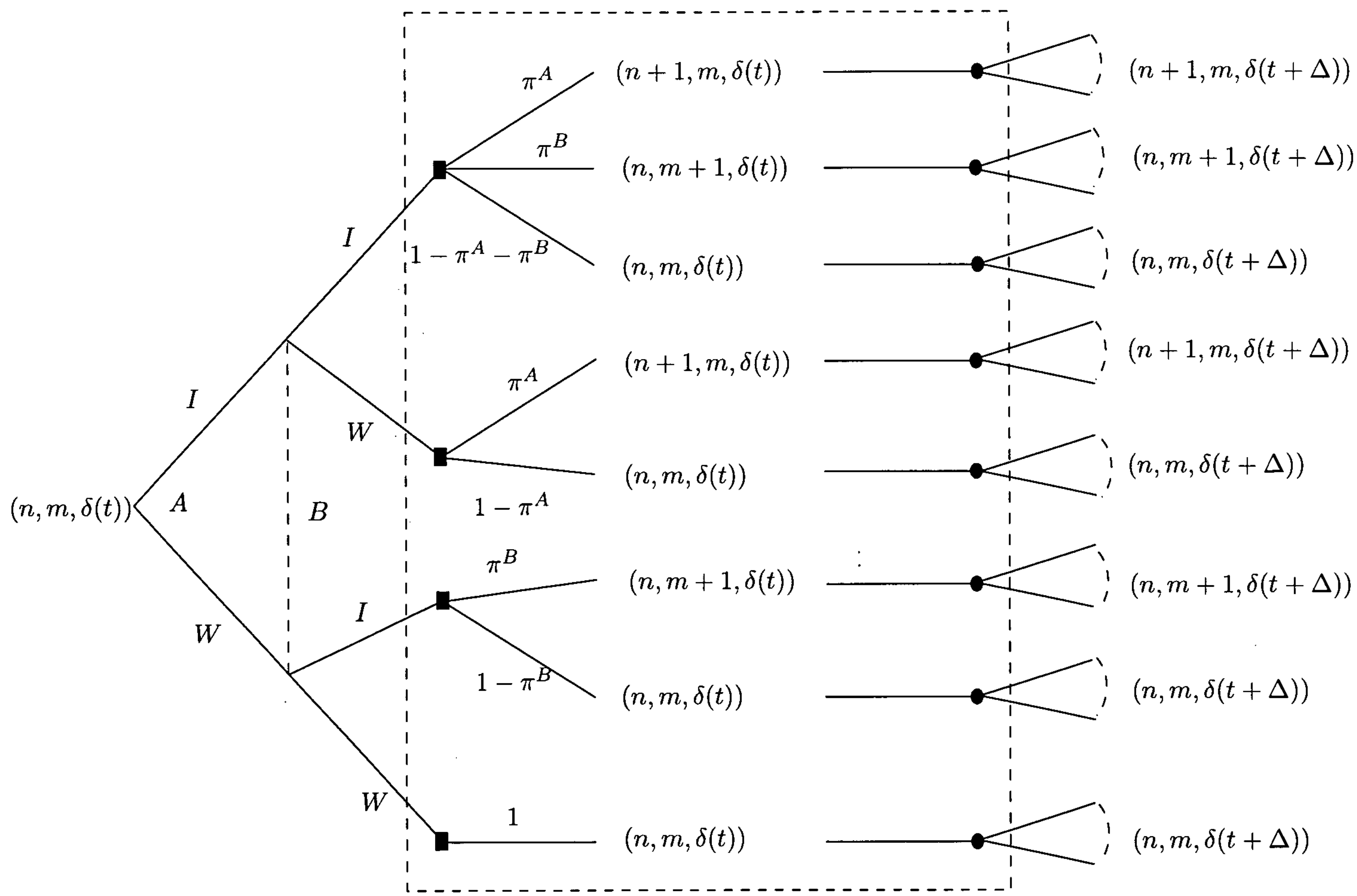

Figure C.1. Graphical representation of the patent race. The tree represents the dynamics of the patent race in subgame $(n, m, \delta(t))$. A and B indicate the two players. In every subgame (n, m, δ), each player simultaneously decides whether to *Invest* (I) or *wait* (W). The decision nodes labeled with a square represent *technical uncertainty*. π^A (π^B) is the probability that firm A (B) is successful when investing. The decision nodes labeled with a circle represent *systematic uncertainty*, i.e., the evolution of the potential cash flow upon completion (δ). The dashed box indicates that the resolution of technical and systematic uncertainty is simultaneous in every subgame. The dashed arcs indicate that $\delta(t+\Delta)$ is drawn from a continuous (lognormal) distribution, as defined in (3.4). Δ is the time interval between two subsequent subgames.

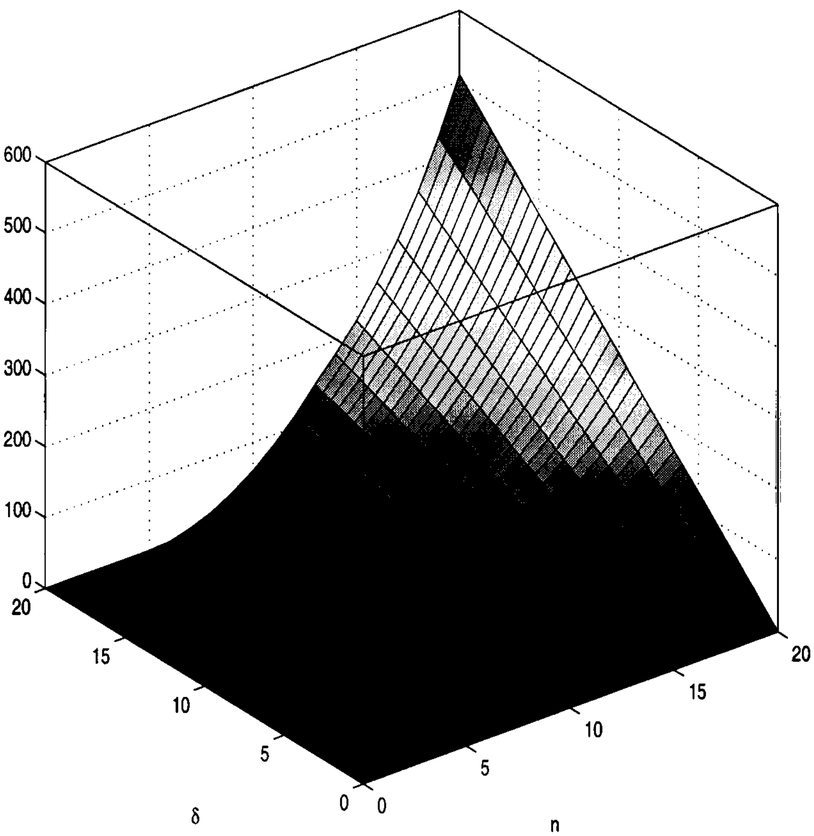

Figure C.2. Family of value functions $V(\cdot, m, \cdot)$ for firm A when B has completed $m = 8$ stages. Values of firm A are plotted against δ, the underlying potential cash flow from the completed project and n, the number of stages completed by firm A. Parameter values are listed in Table III.

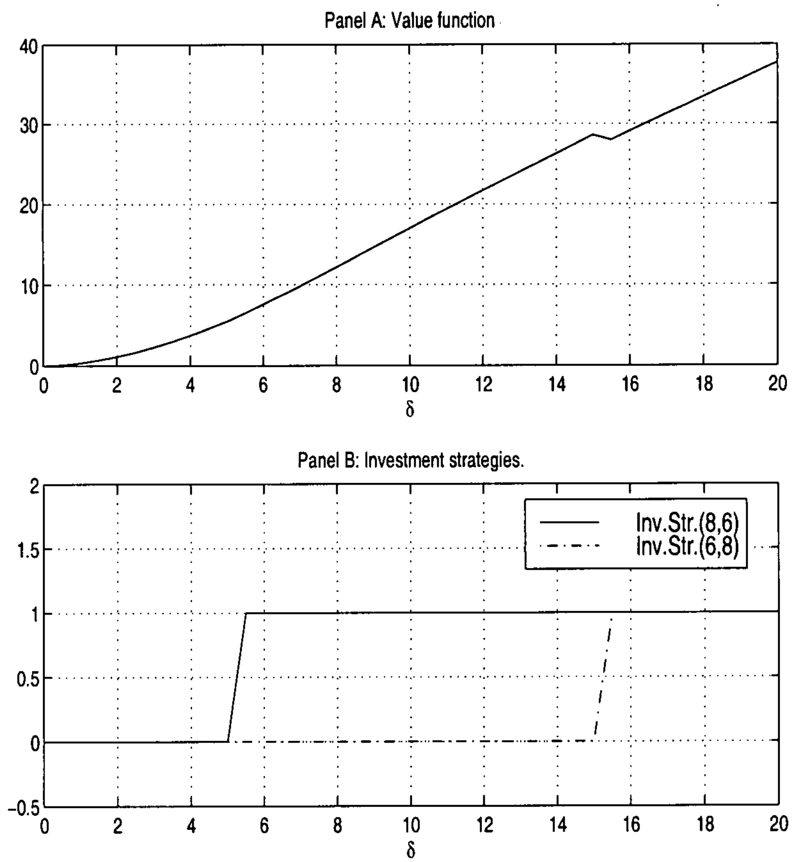

Figure C.3. Discontinuity in the Value Function. Panel A shows the value function (as a function of δ) of firm A in subgame (8,6) and Panel B reports the investment strategies of A and B as a function of δ. In Panel B the solid line indicates the investment strategy fro firm A (leader) and the dotted line is the investment strategy for firm B (follower). '0' indicates that a project is mothballed and '1' indicates that a project is activated. Parameter values are listed in Table III.

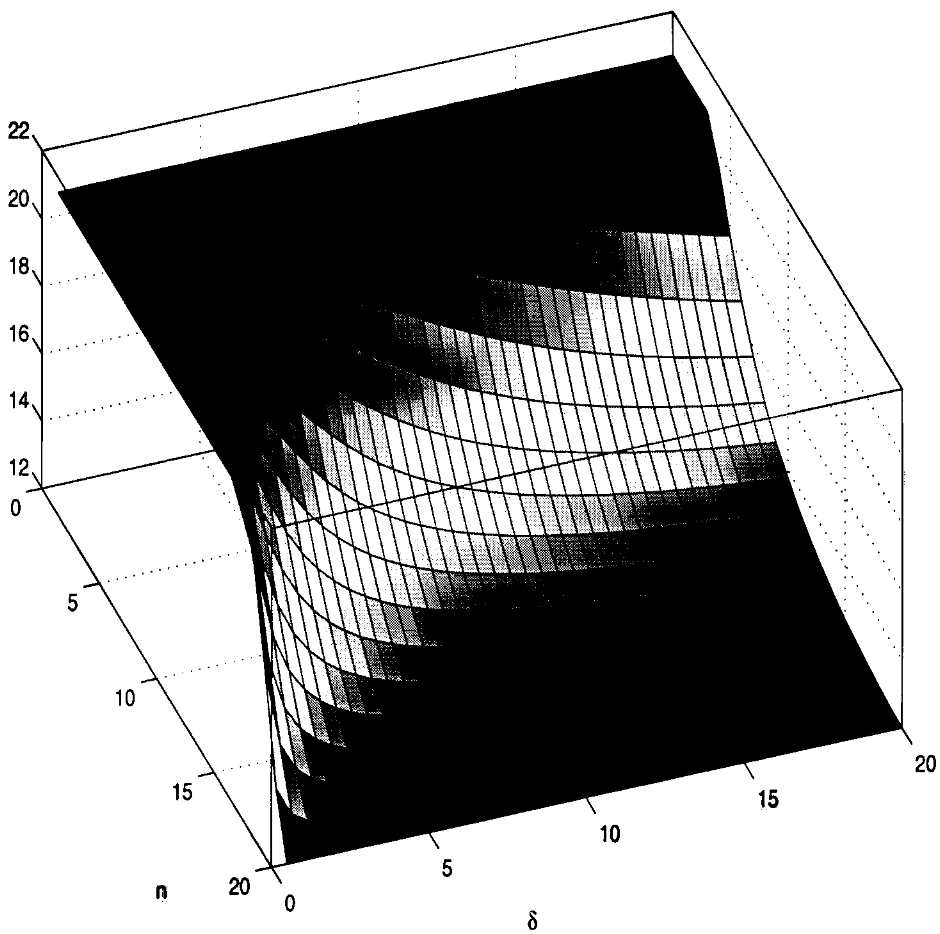

Figure C.4. Risk Premia of Firm A when firm B has completed $m = 1$ stages. The risk premia are expressed as a % p.a. and plotted as a function of δ and of the number of stages (n) completed by firm A. The overall project requires 20 stages to be completed. The risk premium is defined as the difference between the continuously compounded expected return of the project and the riskless rate (see equation (2.25)). Parameter values are listed in Table III.

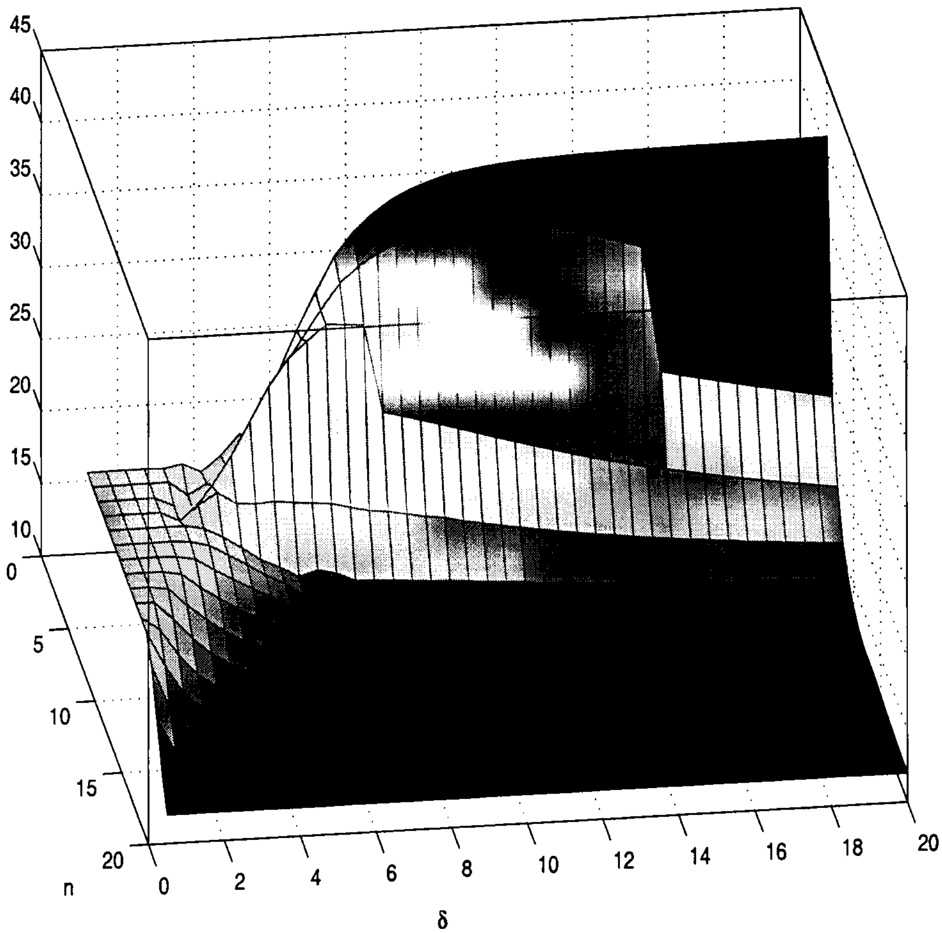

Figure C.5. Risk Premia of Firm A when firm B has completed $m = 8$ stages. The risk premia are expressed as a % p.a. and plotted as a function of δ and of the number of stages (n) completed by firm A. The overall project requires 20 stages to be completed. The risk premium is defined as the difference between the continuously compounded expected return of the project and the riskless rate (see equation (2.25)). Parameter values are listed in Table III.

Table IV. Value functions $V(n, m, \delta)$, for early stages (Panel A), intermediate stages (Panel B) and late stages (Panel C) of developments of an R&D project requiring 20 stages to be completed. $V(n, m)$ represents both the value of Firm A in subgame (n, m) and the value of Firm B in subgame (m, n). Parameter values are listed in Table III.

	Panel A: Early Stages				Panel B: Intermediate Stages				Panel C: Late Stages			
δ	$V(0,0)$	$V(0,1)$	$V(1,0)$	$V(1,1)$	$V(10,10)$	$V(10,11)$	$V(11,10)$	$V(11,11)$	$V(15,15)$	$V(15,16)$	$V(16,15)$	$V(16,16)$
0.0	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
1.0	0.000*	0.000*	0.004*	0.001*	0.283*	0.114*	0.821*	0.417*	2.048	0.842	5.279	3.405
2.0	0.000*	0.000*	0.012*	0.005*	0.936*	0.386*	2.704*	1.222	7.457	4.689	13.908	10.434
3.0	0.001*	0.000*	0.025*	0.010*	1.877	0.809*	5.529	3.081	13.597	9.221	23.067	17.894
4.0	0.002*	0.000*	0.041*	0.016*	3.500	1.813	7.985	5.425	19.933	13.928	32.380	25.455
5.0	0.002*	0.000*	0.060*	0.023*	5.410	3.225	11.128	8.055	26.347	18.702	41.755	33.055
6.0	0.003*	0.000*	0.082*	0.032*	7.508	4.832	14.416	10.853	32.800	23.508	51.158	40.672
7.0	0.004*	0.000*	0.108*	0.042*	9.732	6.565	17.821	13.756	39.274	28.332	60.578	48.298
8.0	0.005*	0.000*	0.135*	0.053*	12.043	8.383	21.307	16.728	45.762	33.166	70.008	55.930
9.0	0.007*	0.000*	0.166*	0.065*	14.414	10.262	24.851	19.749	52.258	38.008	79.445	63.565
10.0	0.008*	0.000*	0.199*	0.078*	16.830	12.184	28.436	22.805	58.761	42.855	88.887	71.202
11.0	0.009*	0.000*	0.235*	0.091*	19.281	14.140	32.052	25.888	65.268	47.705	98.332	78.841
12.0	0.011*	0.000*	0.273*	0.106*	21.758	16.121	35.693	28.992	71.778	52.557	107.780	86.481
13.0	0.013*	0.000*	0.313*	0.122*	24.256	18.123	39.353	32.112	78.290	57.411	117.230	94.121
14.0	0.014*	0.000*	0.356*	0.139*	26.772	20.142	43.028	35.245	84.804	62.266	126.681	101.762
15.0	0.016*	0.000*	0.401*	0.156*	29.302	22.174	46.715	38.389	91.320	67.122	136.134	109.403
16.0	0.018*	0.000*	0.448*	0.175*	31.844	24.217	50.413	41.543	97.838	71.980	145.588	117.045
17.0	0.020*	0.000*	0.497*	0.194*	34.395	26.270	54.120	44.703	104.356	76.838	155.043	124.687
18.0	0.022*	0.000*	0.549*	0.214*	36.956	28.331	57.834	47.870	110.875	81.696	164.499	132.329
19.0	0.024*	0.000*	0.602*	0.235*	39.523	30.399	61.555	51.042	117.394	86.555	173.955	139.971
20.0	0.026*	0.000*	0.658*	0.257*	42.097	32.473	65.281	54.219	123.914	91.414	183.412	147.613

* Firm A Moth-Balls.

Table V. Effect of distance to completion and relative distance between firms on Value of R&D project to firm A. The relative distance is defined as the gap between the number of stages completed by the two firms $(n-m)$. If $(n-m)>0$ firm A is leader. Panel A considers value functions in the early stages of development (15 phases to complete). Panel B considers value functions in the intermediate stage of development (10 phases to complete) and Panel C considers value functions in the late stages of development (5 phases to complete). Entries marked with '–' represent infeasible relative distances. Parameter values are listed in Table III.

	Relative distance between firms $(n-m)$												
δ	20	15	10	5	2	1	0*	−1	−2	−5	−10	−15	−20
Panel A: Distance to completion for Firm $A=15$ $(n=5)$. *Early development.*													
0.0	–	–	–	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	–
5.0	–	–	–	1.504	1.503	1.325	0.440	0.082	0.014	0.000	0.000	0.000	–
10.0	–	–	–	4.974	4.970	4.375	1.486	0.277	0.057	0.000	0.000	0.000	–
15.0	–	–	–	10.079	10.071	8.893	3.577	0.880	0.206	0.000	0.000	0.000	–
20.0	–	–	–	16.323	16.304	13.692	5.990	2.277	0.523	0.000	0.000	0.000	–
Panel B: Distance to completion for Firm $A=10$ $(n=10)$. *Intermediate development.*													
0.0	–	–	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	–	–
5.0	–	–	12.297	12.297	11.263	8.115	5.409	3.225	1.520	0.007	0.000	–	–
10.0	–	–	33.771	33.771	26.689	21.595	16.830	12.184	7.757	0.081	0.000	–	–
15.0	–	–	56.325	56.325	43.092	36.215	29.301	22.174	15.072	0.325	0.000	–	–
20.0	–	–	79.172	79.172	59.924	51.211	42.096	32.473	22.704	0.971	0.000	–	–
Panel C: Distance to completion for Firm $A=5$ $(n=15)$. *Late development.*													
0.0	–	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	–	–	–
5.0	–	48.363	48.363	48.315	38.645	33.050	26.347	18.701	10.835	0.000	–	–	–
10.0	–	104.715	104.715	104.196	83.147	72.322	58.760	42.854	26.087	0.000	–	–	–
15.0	–	161.167	161.167	159.176	127.961	111.808	91.320	67.122	41.458	0.000	–	–	–
20.0	–	217.640	217.640	211.024	172.851	151.345	123.914	91.413	56.853	0.000	–	–	–

* Neck-and-neck.

Table VI. Effect of Volatility of cash flows on firm value. Values are computed in three different stages of the development process, initial stages, intermediate stages and late stages. Panel A considers values for low volatility of cash flows ($\sigma_\delta = 0.10$). Panel B considers values for medium volatility levels ($\sigma_\delta = 0.15$) and Panel C for high volatility level ($\sigma_\delta = 0.30$).Parameter values are listed in Table III.

	Early Stages				Intermediate Stages				Late Stages			
δ	$V(0,0)$	$V(0,1)$	$V(1,0)$	$V(1,1)$	$V(10,10)$	$V(10,11)$	$V(11,10)$	$V(11,11)$	$V(15,15)$	$V(15,16)$	$V(16,15)$	$V(16,16)$
Panel A: Low Volatility of Cash Flows ($\sigma_\delta = 0.10$ or 20% p.a.).												
0.0	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
5.0	87.480	73.482	108.203	93.577	159.735	128.942	200.113	168.407	205.696	149.596	273.163	215.552
10.0	199.975	171.392	240.856	211.360	333.832	271.474	413.789	349.824	418.513	305.412	552.546	436.664
15.0	313.037	269.838	374.067	329.626	507.983	414.058	627.510	531.277	631.331	461.228	831.930	657.777
20.0	426.490	368.658	507.655	448.246	682.205	556.706	841.287	712.781	844.152	617.047	1111.316	878.890
Panel B: Medium Volatility of Cash Flows ($\sigma_\delta = 0.15$ or 30% p.a.).												
0.0	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
5.0	27.017	20.886	37.916	30.994	82.384	65.754	107.396	89.910	123.754	89.939	168.297	133.073
10.0	73.552	60.962	95.077	81.606	178.700	144.709	227.949	192.545	254.628	186.098	342.813	271.706
15.0	121.744	102.618	153.835	133.629	275.170	223.803	348.644	295.283	385.503	282.258	517.329	410.340
20.0	170.418	144.733	213.051	186.063	371.698	302.949	469.396	398.063	516.380	378.418	691.847	548.974
Panel C: High Volatility of Cash Flows ($\sigma_\delta = 0.30$ or 60% p.a.).												
0.0	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
5.0	0.002*	0.000*	0.060*	0.023*	5.410	3.225	11.128	8.055	26.347	18.702	41.755	33.055
10.0	0.008*	0.000*	0.199*	0.078*	16.830	12.184	28.436	22.805	58.761	42.855	88.887	71.202
15.0	0.016*	0.000*	0.401*	0.156*	29.302	22.174	46.715	38.389	91.320	67.122	136.134	109.403
20.0	0.026*	0.000*	0.658*	0.257*	42.097	32.473	65.281	54.219	123.914	91.414	183.412	147.613

* Firm *A* Moth-Balls.

Table VII. Value-dissipating effect of R&D competition. The table reports the percentage erosion in value (Φ), computed according to Definition 2.5.1. Panel A contains percentage erosions for early stages of development, Panel B for intermediate stages of development and Panel C for late stages of development. The project requires 20 stage to be completed. Parameter values are listed in Table III.

	Panel A: Early Stages				Panel B: Intermediate Stages				Panel C: Late Stages			
δ	$\Phi(0,0)$	$\Phi(0,1)$	$\Phi(1,1)$	$\Phi(2,2)$	$\Phi(10,10)$	$\Phi(10,11)$	$\Phi(11,11)$	$\Phi(12,12)$	$\Phi(15,15)$	$\Phi(15,16)$	$\Phi(16,16)$	$\Phi(17,17)$
1.0	4.153	0.000	25.649	32.193	30.427	15.821	29.089	28.415	20.894	16.999	13.429	7.8049
2.0	3.969	0.000	25.637	32.193	30.209	15.815	37.041	25.649	8.832	8.132	5.332	2.4422
3.0	3.992	0.000	25.640	32.190	30.479	14.652	20.918	15.730	5.102	4.486	2.872	1.1375
4.0	4.040	0.000	25.642	32.191	21.111	18.942	16.034	11.654	3.427	3.058	1.839	0.6491
5.0	4.024	0.000	25.640	32.192	17.608	16.145	12.883	9.092	2.520	2.248	1.294	0.4095
6.0	4.014	0.000	25.639	32.192	14.983	13.875	10.662	7.402	1.962	1.741	0.966	0.2740
7.0	4.002	0.000	25.640	32.192	12.969	11.996	9.063	6.232	1.590	1.405	0.753	0.1909
8.0	4.011	0.000	25.640	32.189	11.401	10.473	7.878	5.382	1.327	1.171	0.608	0.1372
9.0	4.010	0.000	25.640	32.191	10.168	9.242	6.974	4.739	1.134	1.001	0.503	0.1011
10.0	4.000	0.000	25.640	32.193	9.183	8.287	6.264	4.237	0.987	0.873	0.426	0.0761
11.0	4.007	0.000	25.640	32.192	8.382	7.616	5.694	3.834	0.873	0.774	0.367	0.0583
12.0	4.011	0.000	25.640	32.192	7.722	7.051	5.227	3.505	0.781	0.694	0.322	0.0454
13.0	4.007	0.000	25.640	32.183	7.170	6.570	4.836	3.232	0.706	0.630	0.285	0.0359
14.0	4.008	0.000	25.640	32.194	6.700	6.155	4.505	3.001	0.644	0.576	0.256	0.0288
15.0	4.007	0.000	25.638	32.235	6.297	5.795	4.222	2.804	0.592	0.531	0.231	0.0233
16.0	4.010	0.000	25.638	32.255	5.947	5.479	3.976	2.635	0.548	0.492	0.211	0.0191
17.0	4.007	0.000	25.638	32.301	5.640	5.200	3.762	2.487	0.510	0.459	0.194	0.0158
18.0	4.009	0.000	25.639	32.184	5.368	4.952	3.573	2.358	0.477	0.430	0.179	0.0132
19.0	4.010	0.000	25.641	32.241	5.127	4.731	3.405	2.243	0.448	0.405	0.166	0.0111
20.0	4.009	0.000	25.643	32.092	4.911	4.532	3.256	2.141	0.422	0.382	0.155	0.0094

Table VIII. Impact of Preemption on Investment Decisions. Panel A reports the value function V and investment strategies when firm A has completed $n = 8$ stages and faces a competitor far behind ("Quasi-monopoly" scenario). It also reports the value function V^J and investment strategy of the RJV. The last column contains the % erosion as defined in Definition 2.5.1. Panel B contains the same information for the case of a "Neck-and-Neck" scenario. Parameter values are listed in Table III.

	Panel A: "Quasi-Monopoly"					Panel B: "Neck-and-Neck"				
δ	$V(8,0)$	*Invest?*	$V^J(8,0)$	*Invest?*	$\Phi(8,0)$	$V(8,8)$	*Invest?*	$V^J(8,8)$	*Invest?*	$\Phi(8,8)$
0.0	0.00	No	0.00	No	0	0.00	No	0.00	No	0.00
1.0	0.35	No	0.35	No	0	0.12	No	0.37	No	35.11
2.0	1.17	No	1.17	No	0	0.40	No	1.24	No	35.13
3.0	2.37	No	2.37	No	0	0.81	No	2.49	No	34.98
4.0	3.89	No	3.89	No	0	1.17	Yes	4.10	No	42.60
5.0	5.72	No	5.72	No	0	2.02	Yes	6.03	No	32.68
6.0	7.87	Yes	7.87	Yes	0	3.01	Yes	8.24	No	26.71
7.0	10.30	Yes	10.30	Yes	0	4.12	Yes	10.84	Yes	23.86
8.0	12.87	Yes	12.87	Yes	0	5.33	Yes	13.67	Yes	22.04
9.0	15.55	Yes	15.55	Yes	0	6.61	Yes	16.61	Yes	20.40
10.0	18.30	Yes	18.30	Yes	0	7.95	Yes	19.63	Yes	18.96
11.0	21.10	Yes	21.10	Yes	0	9.34	Yes	22.71	Yes	17.71
12.0	23.95	Yes	23.95	Yes	0	10.77	Yes	25.84	Yes	16.62
13.0	26.83	Yes	26.83	Yes	0	12.23	Yes	29.00	Yes	15.65
14.0	29.74	Yes	29.74	Yes	0	13.71	Yes	32.20	Yes	14.80
15.0	32.67	Yes	32.67	Yes	0	15.22	Yes	35.42	Yes	14.04
16.0	35.62	Yes	35.62	Yes	0	16.74	Yes	38.66	Yes	13.37
17.0	38.58	Yes	38.58	Yes	0	18.28	Yes	41.93	Yes	12.77
18.0	41.56	Yes	41.56	Yes	0	19.83	Yes	45.21	Yes	12.23
19.0	44.55	Yes	44.55	Yes	0	21.40	Yes	48.50	Yes	11.75
20.0	47.55	Yes	47.55	Yes	0	22.97	Yes	51.81	Yes	11.31

Table IX. Risk premia $RP(n, m)$ earned on an ownership claim to venture A. Risk premia are expressed as % p.a. and defined in equation (2.25). Panel A reports risk premia in early stages of development, Panel B reports risk premia in intermediate stages of development and Panel C reports risk premia in late stages of development. Parameter values are listed in Table III.

	Panel A: $RP(n,m)$, Early Stages				Panel B: $RP(n,m)$, Intermediate Stages				Panel C: $RP(n,m)$, Late Stages			
δ	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$	$(10,10)$	$(10,11)$	$(11,10)$	$(11,11)$	$(15,15)$	$(15,16)$	$(16,15)$	$(16,16)$
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.0	20.67	0.00	20.67	20.67	20.71	20.58	20.67	20.52	21.08	25.78	17.28	19.26
2.0	20.69	0.00	20.69	20.69	21.93	24.17	19.92	20.82	17.78	19.85	14.90	16.04
3.0	20.69	0.00	20.69	20.69	21.59	29.89	18.40	21.59	15.90	17.05	14.10	14.66
4.0	20.70	0.00	20.70	20.70	20.94	25.38	17.18	20.02	14.92	15.68	13.61	13.97
5.0	20.70	0.00	20.70	20.70	20.03	23.61	16.41	18.85	14.32	14.89	13.30	13.56
6.0	20.70	0.00	20.70	20.70	19.15	21.95	16.05	17.88	13.93	14.37	13.09	13.29
7.0	20.70	0.00	20.70	20.70	18.37	20.61	15.70	17.13	13.64	14.01	12.94	13.10
8.0	20.70	0.00	20.70	20.70	17.68	19.54	15.38	16.53	13.43	13.74	12.82	12.96
9.0	20.70	0.00	20.70	20.70	17.12	18.70	15.10	16.06	13.27	13.54	12.73	12.85
10.0	20.70	0.00	20.70	20.70	16.65	18.02	14.85	15.67	13.14	13.37	12.66	12.76
11.0	20.70	0.00	20.70	20.70	16.26	17.46	14.63	15.35	13.03	13.24	12.60	12.69
12.0	20.70	0.00	20.70	20.70	15.93	17.00	14.45	15.08	12.94	13.13	12.55	12.63
13.0	20.70	0.00	20.70	20.70	15.65	16.60	14.28	14.85	12.87	13.04	12.51	12.58
14.0	20.70	0.00	20.70	20.70	15.40	16.27	14.14	14.66	12.81	12.96	12.47	12.53
15.0	20.70	0.00	20.70	20.70	15.18	15.98	14.01	14.48	12.75	12.90	12.44	12.50
16.0	20.70	0.00	20.70	20.70	14.99	15.73	13.90	14.33	12.70	12.84	12.41	12.47
17.0	20.70	0.00	20.70	20.70	14.83	15.50	13.80	14.20	12.66	12.79	12.39	12.44
18.0	20.70	0.00	20.70	20.70	14.68	15.31	13.71	14.08	12.62	12.74	12.36	12.41
19.0	20.70	0.00	20.70	20.70	14.54	15.13	13.63	13.97	12.59	12.70	12.35	12.39
20.0	20.70	0.00	20.70	20.70	14.42	14.97	13.55	13.87	12.56	12.66	12.33	12.37

Table X. Effect of distance to completion and relative distance between firms on Risk Premia of firm A. The relative distance is defined as the gap between the number of stages completed by the two firms $(n-m)$. If $n-m>0$ firm A is leader. Panel A considers risk premia in the early stages of development (15 phases to complete). Panel B considers risk premia in the intermediate stage of development (10 phases to complete) and Panel C considers risk premia in the late stages of development (5 phases to complete). Entries marked with '–' represent infeasible relative distances. Risk premia are defined in equation (2.25). Risk premia are set equal to zero when the value function is zero.Parameter values are listed in Table III.

	Relative distance between firms $(n-m)$												
δ	20	15	10	5	2	1	0*	−1	−2	−5	−10	−15	−20
Panel A: Distance to completion for Firm $A=15$ $(n=5)$. *Early development.*													
0.0	–	–	–	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	–
5.0	–	–	–	20.700	20.700	20.700	20.729	20.472	21.141	0.000	0.000	0.000	–
10.0	–	–	–	20.669	20.668	20.394	20.469	32.254	35.966	0.000	0.000	0.000	–
15.0	–	–	–	19.455	19.447	17.679	18.245	34.139	38.971	0.000	0.000	0.000	–
20.0	–	–	–	18.163	18.141	15.556	17.913	23.102	38.095	0.000	0.000	0.000	–
Panel B: Distance to completion for Firm $A=10$ $(n=10)$. *Intermediate development.*													
0.0	–	–	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	–	–
5.0	–	–	18.501	18.501	17.225	16.885	20.031	23.613	26.974	41.478	0.000	–	–
10.0	–	–	15.490	15.490	14.329	15.389	16.656	18.021	19.787	41.629	0.000	–	–
15.0	–	–	14.341	14.341	13.683	14.489	15.189	15.982	17.058	39.388	0.000	–	–
20.0	–	–	13.758	13.758	13.436	13.948	14.422	14.975	15.738	22.173	0.000	–	–
Panel C: Distance to completion for Firm $A=5$ $(n=15)$. *Late development.*													
0.0	–	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	–	–	–
5.0	–	13.758	13.758	13.731	13.359	13.727	14.326	14.890	15.792	0.000	–	–	–
10.0	–	12.852	12.852	12.700	12.817	12.973	13.142	13.379	13.786	0.000	–	–	–
15.0	–	12.560	12.560	12.272	12.559	12.652	12.755	12.901	13.161	0.000	–	–	–
20.0	–	12.417	12.417	12.210	12.424	12.489	12.563	12.668	12.857	0.000	–	–	–

* Neck-and-neck.

Table XI. Risk Premia (% p.a.) for a value-weighted portfolio of R&D firms and for a RJV. Risk premia RP^{pf} are derived in Definition 2.5.5 and RP^J are derived in Definition 2.5.6. Panel A reports the relationship between RP^{pf} and RP^J in a "Quasi-monopoly" scenario for different stages of development (early, intermediate and late development). Panel B does the same for the case of "Neck-and-Neck"scenarios. The projects requires 20 stages to be completed. Parameter values are listed in Table III.

Panel A: "Quasi-Monopoly"

	Early Stage			Intermediate Stage			Late Stage		
δ	$RP^{pf}(5,0)$	$RP^J(5,0)$	$RP^{pf}-RP^J$	$RP^{pf}(10,0)$	$RP^J(10,0)$	$RP^{pf}-RP^J$	$RP^{pf}(15,0)$	$RP^J(15,0)$	$RP^{pf}-RP^J$
2.0	20.70*	20.70*	0.00	20.70*	20.70*	0.00	16.56	16.56	0.00
4.0	20.70*	20.70*	0.00	19.51	19.51	0.00	14.22	14.22	0.00
6.0	20.70*	20.70*	0.00	17.61	17.61	0.00	13.45	13.45	0.00
8.0	20.70*	20.70*	0.00	16.32	16.32	0.00	13.07	13.07	0.00
10.0	20.66*	20.66*	0.00	15.49	15.49	0.00	12.85	12.85	0.00
12.0	20.49*	20.49*	0.00	14.92	14.92	0.00	12.70	12.70	0.00
14.0	19.71	19.71	0.00	14.50	14.50	0.00	12.60	12.60	0.00
16.0	19.19	19.19	0.00	14.19	14.19	0.00	12.52	12.52	0.00
18.0	18.66	18.66	0.00	13.95	13.95	0.00	12.46	12.46	0.00
20.0	18.16	18.16	0.00	13.75	13.75	0.00	12.41	12.41	0.00

Panel B: "Neck-and-Neck"

	Early Stage			Intermediate Stage			Late Stage		
δ	$RP^{pf}(5,5)$	$RP^J(5,5)$	$RP^{pf}-RP^J$	$RP^{pf}(10,10)$	$RP^J(10,10)$	$RP^{pf}-RP^J$	$RP^{pf}(15,15)$	$RP^J(15,15)$	$RP^{pf}-RP^J$
2.0	20.70*	20.70*	0.00	21.46*	20.70*	0.76	17.78	16.50	1.28
4.0	20.66*	20.70*	-0.04	20.88	19.42*	1.45	14.92	14.38	0.54
6.0	21.10*	21.59*	-0.49	19.15	17.60	1.54	13.93	13.62	0.31
8.0	22.98*	23.83*	-0.85	17.68	16.39	1.28	13.43	13.23	0.20
10.0	20.46	26.20*	-5.74	16.65	15.62	1.03	13.14	12.99	0.14
12.0	19.73	20.61*	-0.88	15.93	15.08	0.84	12.94	12.83	0.11
14.0	18.89	20.37*	-1.48	15.40	14.69	0.71	12.81	12.71	0.09
16.0	18.47	18.95	-0.48	14.99	14.38	0.61	12.70	12.63	0.07
18.0	18.19	18.42	-0.23	14.68	14.14	0.54	12.62	12.56	0.06
20.0	17.96	17.92	0.03	14.42	13.94	0.48	12.56	12.50	0.05

* Both Units Moth-balled ($[W, W]$).

Appendix D

Feasibility Set for the One-Asset case

In this Appendix we gain some intuition on the properties of the feasibility set by looking at a simple one-asset problem. We are interested in characterizing the set of feasible trades available at a generic time t. Simplifying Definition 3.2.1 to handle the single-asset case, we deduce that the trades feasible at time t are represented by the set of points $(\Delta Y, \Delta X)$ such that

$$\begin{aligned}
\Delta Y &\leq -\Delta X - \tau_g G(t) && \text{(D.1)}\\
\Delta X &\geq -X(t) && \text{(D.2)}\\
\Delta Y &\geq \max_{\tilde{R}}\{-Y(t) - \tilde{R}X - \tilde{R}\Delta X\} && \text{(D.3)}
\end{aligned}$$

Equation (D.1) is the self financing constraint, equation (D.2) is the short-selling constraint while (D.3) follows from the requirement of non-negativity of liquidating wealth. In equation (D.3) the $\max(\cdot)$ is taken over all possible realizations of the return on the asset $X(t)$. In the above equations, the stock $X(t)$, the bank account $Y(t)$ and the basis $B(t)$ evolve as follows

$$\begin{aligned}
X(t+1) &= (X(t) + \Delta X)\tilde{R} && \text{(D.4)}\\
Y(t+1) &= Y(t) - \Delta X - \tau G(t) && \text{(D.5)}\\
B(t+1) &= \begin{cases} B(t) + \Delta(t)\left(1 + \chi_{\{\Delta(t)<0\}}\left(\frac{B(t)}{X(t)} - 1\right)\right) & \text{if} \quad B(t) < X(t) \\ X(t) + \Delta(t) & \text{if} \quad B(t) > X(t) \end{cases} && \text{(D.6)}
\end{aligned}$$

while the gains $G(t)$ are

$$G(t) = \begin{cases} -\Delta X\left(1 - \frac{B(t)}{X(t)}\right)\chi_{\{\Delta X<0\}} & \text{if} \quad B(t) < X(t) \\ X(t) - B(t) & \text{if} \quad B(t) > X(t) \end{cases} \tag{D.7}$$

Panel A: $B(t) < X(t)$ **Panel B:** $B(t) > X(t)$

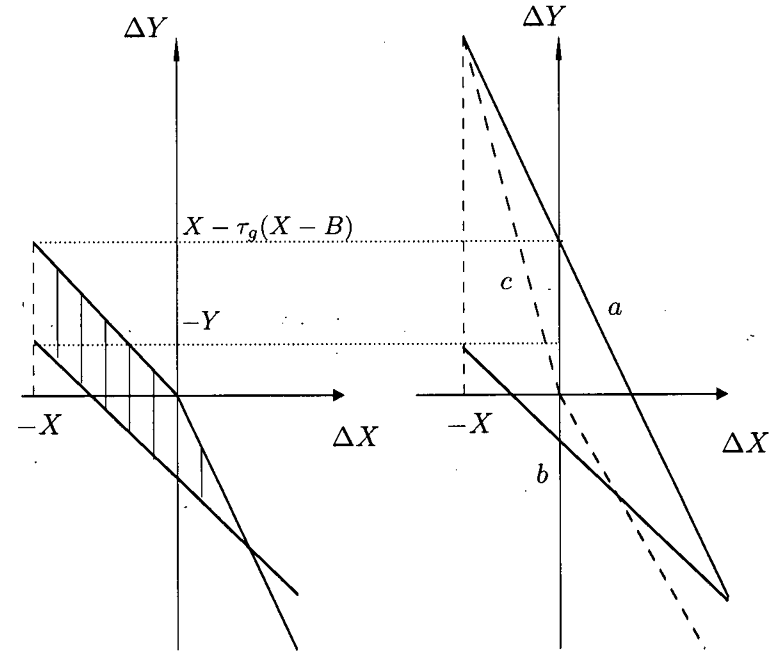

Figure D.1. Feasibility Set at time 0. This figures displays the feasibility set at time $t = 0$ for a one-asset portfolio problem with capital gains taxes. Panel A refers to the case of an embedded capital gain while Panel B refers to the case of an embedded capital loss $B(t) > X(t)$. The shaded region in Panel A is the feasibility set described by equations (D.1) and (D.3). Line a is the self financing constraint (D.1) when wash sales are allowed while the dashed line c is the self financing constraint (D.3) when wash sales are not allowed. Line b is the non-negativity of liquidating wealth requirement (D.3).

Figure D.1 represents the feasibility set described by equations (D.1) and (D.3). Note that when there is an embedded gain (Panel A) the feasibility set has a "kink" at $\Delta X = 0$. This kink is responsible for the no-trade region we observe in the paper. The requirement of non-negative liquidating wealth $(X(t) + Y(t) - \tau_g(X(t) - B(t)) \geq 0)$ implies that the trade $\Delta X = -X$ is always feasible, guaranteeing the non-emptiness of the feasibility set. It is important to realize that if we only impose non-negativity of *paper* wealth, i.e., $X(t) + Y(t) \geq 0$ the feasibility set might be empty for some initial states. Panel B shows the feasibility sets when there is an embedded loss. The solid line a is the self financing constraint when wash sales are allowed. Notice that wash sales implies a positive parallel shift of the no-tax self financing constraint for the amount of the tax rebate $-\tau_g(X - B) > 0$. If we do not allow wash sales, equations (D.6) and (D.7) become

$$B(t+1) = B + \Delta X \left(1 + \chi_{\{\Delta X < 0\}} \left(\frac{B(t)}{X(t)}\right)\right) \tag{D.8}$$

and

$$G(t) = -\Delta X \chi_{\{\Delta X < 0\}} \left(1 - \frac{B(t)}{X(t)}\right) \tag{D.9}$$

In this case the self financing constraint has a "kink" in the opposite direction than the one occurring in the gain region (see the dashed line b in Panel B). When wash sales are not allowed, the feasibility set in the loss region may not be a non-convex set as Figure D.1 illustrates.

Appendix E

Proofs of Propositions in Chapter 3

Proof of Proposition 3.2.2 (Compactness of Feasibility Set)

For simplicity and without loss of generality we assume that the two assets do not pay dividends. The proof for the case of dividends follows identical steps.

Let us first consider a one-period problem. Let

$$\delta_1 = (\Delta_0 Y, \mathbf{\Delta}_0 \mathbf{X}) = (\Delta_0 Y, \Delta_0 X_1, \Delta_0 X_2)$$

be a generic investment strategy. Let $\tilde{\boldsymbol{R}} = (\tilde{R}_1, \tilde{R}_2)$ be a K-variate random vector with realizations $(\tilde{R}_1^{(k)}, \tilde{R}_2^{(k)})$, $k = 1, \ldots, K$, representing the (ex-dividend) return on the two assets.[1] Since the bank account earns zero interest, a necessary condition for no arbitrage is that

$$r_i = \min\{\tilde{R}_i^{(k)}, k = 1, \ldots, K\} < 1, \quad i = 1, 2. \tag{E.1}$$

Let $\mathbf{X}(t) = (X_1(t), X_2(t))$ and $\mathbf{B}(t) = (B_1(t), B_2(t))$ be the vector of asset holding and bases at a generic time t and

$$\mathbf{Z}(0) = (Y(0), \mathbf{X}(0), \mathbf{B}(0)), \quad (Y(0), \mathbf{X}(0)) \in \mathcal{S}_{\mathbf{B}(0)}$$

an initial feasible state.[2] The set $\Phi_1(\mathbf{Z}(0))$ of feasible trades $(\Delta_0 Y, \mathbf{\Delta}_0 \mathbf{X})$ from state $\mathbf{Z}(0)$ is defined as follows

$$\Phi_1(\mathbf{Z}(0)) = \{(\Delta_0 Y, \mathbf{\Delta}_0 \mathbf{X}) \in \mathbb{R}^3 : \Delta_0 Y \leq f_1(\mathbf{\Delta}_0 \mathbf{X}, \mathbf{Z}(0)), \Delta_0 Y \geq f_2(\mathbf{\Delta}_0 \mathbf{X}, \mathbf{Z}(0), \tilde{\boldsymbol{R}}), \mathbf{\Delta}_0 \mathbf{X} \geq -\mathbf{X}(0)\} \tag{E.2}$$

[1] In the model $(\tilde{R}_1, \tilde{R}_2)$ follows a trinomial distribution, hence $k = 3$.

[2] See Definition 3.2.1.

where the functions $f_1(\cdot)$ and $f_2(\cdot)$ are directly obtained from Definition 3.2.1

$$f_1(\mathbf{\Delta_0 X}, \mathbf{Z}(0)) = -\sum_{i=1}^{2} \Delta_0 X_i - \tau_g \sum_{i=1}^{2} G_i(\Delta_0 X_i, \mathbf{Z}(0)) \tag{E.3}$$

$$f_2(\mathbf{\Delta_0 X}, \mathbf{Z}(0)) = -Y(0) - \sum_{i=1}^{2} r_i(X_i(0) + \Delta_0 X_i) \tag{E.4}$$

and where the gain $G_i(\Delta_0 X_i, \mathbf{Z}(0))$ is defined as (see equation (3.5))

$$G_i(\Delta_0 X_i, \mathbf{Z}(0)) = \begin{cases} -\Delta_0 X_i(t)\left(1 - \frac{B_i(0)}{X_i(0)}\right)\chi_{\{\Delta_0 X_i < 0\}} & \text{if } B_i(0) < X_i(0) \\ X_i(0) - B_i(0) & \text{if } B_i(0) > X_i(0) \end{cases}, \quad i = 1, 2. \tag{E.5}$$

In (E.4) the presence of r_i (see equation (E.1)) is due to the fact that feasibility needs to be guaranteed for all possible realizations of $\tilde{R}_i$. Notice that at $\mathbf{\Delta_0 X} = (-X_1(0), -X_2(0))$

$$f_1(-\mathbf{X}_0) = -\sum_{i=1}^{2}(X_i(0) - \tau_g(X_i(0) - B_i(0))). \tag{E.6}$$

$$f_2(-\mathbf{X}_0) = -Y(0). \tag{E.7}$$

Since $\mathbf{Z}(0) \in \mathcal{S}_{\mathbf{B}(0)}$, $f_1(-\mathbf{X}_0) \geq f_2(-\mathbf{X}_0)$ and therefore $\Phi_1(\mathbf{Z}(0))$ is non-empty. Equations (E.3), (E.4) and (E.5) imply that $f_1(\cdot)$ and $f_2(\cdot)$ are continuous function of $\mathbf{\Delta_0 X}$.[3] Therefore the set $\Phi_1(\mathbf{Z}(0))$ is closed. Finally, notice that

$$\lim_{\Delta_0 X_i \to \infty} \frac{\partial f_1(\mathbf{\Delta_0 X})}{\Delta_0 X_i} = -1 < -r_i = \lim_{\Delta_0 X_i \to \infty} \frac{\partial f_2(\mathbf{\Delta}_0 X)}{\Delta_0 X_i}$$

Hence, since $\mathbf{\Delta_0 X} \geq -\mathbf{X}(0)$ and $f_1(-\mathbf{X}(0)) \geq f_2(-\mathbf{X}(0))$, for every $\Delta_0 X_i$ $i = 1, 2$, there exists a $\Delta_0^* X_{3-i}(\Delta_0 X_i)$, $-X_{3-i}(0) \leq \Delta_0^* X_{3-i}(\Delta_0 X_i) < \infty$, such that

$$f_1(\Delta_0 X_1, \Delta_0 X_2) \geq f_2(\Delta_0 X_1, \Delta_0 X_2), \text{ for } -X_i(0) \leq \Delta_0 X_i \leq \Delta_0^* X_i(\Delta_0 X_{3-i})$$

$$f_1(\Delta_0 X_1, \Delta_0 X_2) < f_2(\Delta_0 X_1, \Delta_0 X_2), \text{ for } \Delta_0 X_i > \Delta_0^* X_i(\Delta_0 X_{3-i})$$

which is sufficient to guarantee that $\Phi_1(\mathbf{Z}(0))$ is bounded. We hence proved that $\Phi_1(\mathbf{Z}(0))$ is a non-empty, closed and bounded subset of $\mathbb{R}^3$ and therefore $\Phi_1(\mathbf{Z}(0))$ is compact.

Let us now consider a $T = 2$ period problem. Let $\{\mathcal{F}\} = \{\mathcal{F}_0, \mathcal{F}_1\}$ be the sequence of filtrations generated by $\tilde{\boldsymbol{R}}$. Since $\tilde{\boldsymbol{R}}$ is a K-variate random variable, the number of element of $\mathcal{F}_1$ is $N_1 = K$. Trivially, $N_0 = 1$. Let

$$\delta_2 = (\mathbf{\Delta}_0^{N_0}, \mathbf{\Delta}_1^{N_1}) \in \mathbb{R}^{3\times(N_0+N_1)}$$

[3]More precisely, $f_1(\cdot)$ can be thought of as the product of two piecewise linear functions $g(\Delta_0 X_1)$ and $h(\Delta_0 X_2)$, each with a "kink" in $\Delta_0 X_i = 0$, $i = 1, 2$. See Appendix D.

be a generic investment strategy in the two-period problem.[4] The set $\Phi_2(\mathbf{Z}(0))$ of feasible trades $(\boldsymbol{\Delta}_0^{N_0}, \boldsymbol{\Delta}_1^{N_1})$ from state $\mathbf{Z}(0)$ can be defined as follows

$$\Phi_2(\mathbf{Z}(0)) = \left\{(\boldsymbol{\Delta}_0^{N_0}, \boldsymbol{\Delta}_1^{N_1}) \in \mathbb{R}^{3\times(N_0+N_1)} : \boldsymbol{\Delta}_0 \in \Phi_1(\mathbf{Z}(0)), \boldsymbol{\Delta}_1^{N_1} \in \Gamma_2(\boldsymbol{\Delta}_0^{N_0})\right\}$$

where $\boldsymbol{\Delta}_1^{N_1} \in \Gamma_2(\boldsymbol{\Delta}_0^{N_0})$ if and only if for $k = 1, \ldots, N_1$

$$\Delta_1^{(k)} Y \leq -\Delta_1^{(k)} X - \tau_g \sum_{i=1}^{2} G_i^{(k)}(\Delta_i^{(k)} X_i, \mathbf{Z}^{(k)}(\boldsymbol{\Delta}_0)) \tag{E.8}$$

$$\Delta_1^{(k)} Y(\Delta) \geq -Y^{(k)} - \sum_{i=1}^{2} r_i(X_i^{(k)} - \Delta^{(k)} X_i) \tag{E.9}$$

$$\Delta^{(k)} X_i \geq -X_i^{(k)}(\boldsymbol{\Delta}_0), \quad i = 1, 2. \tag{E.10}$$

Equations (E.8) to (E.10) are the equivalent at time $t = 1$ of the conditions in (E.2) defining $\Phi_1(\mathbf{Z}(0))$. In the above equations, $\Delta^{(k)} X_i$ refers to the trade in asset i, $i = 1, 2$ in the k^{th} element of $\mathcal{F}_1$,

$$\mathbf{Z}^{(k)}(\boldsymbol{\Delta}_0) = [\mathbf{X}^{(k)}(\boldsymbol{\Delta}_0), Y^{(k)}(\boldsymbol{\Delta}_0), \mathbf{B}^{(k)}(\boldsymbol{\Delta}_0)], \tag{E.11}$$

$$X_i^{(k)}(\boldsymbol{\Delta}_0) = (X_i(0) + \Delta_0 X)\tilde{R}_i^{(k)}, \quad i = 1, 2 \tag{E.12}$$

$$Y^{(k)}(\boldsymbol{\Delta}_0) = Y(0) - \sum_{i=1}^{2} \Delta_0 X_i - \tau_g \sum_{i=1}^{2} G_i(\Delta_0 X_i, \mathbf{Z}(0)) \tag{E.13}$$

$$B_i^{(k)}(\boldsymbol{\Delta}_0) = \begin{cases} B_i(0) + \Delta_0 X_i \left(1 + \chi_{\{\Delta_0 X_i < 0\}} \left(\frac{B(0)}{X_i(0)} - 1\right)\right) & \text{if } \frac{B_i(0)}{X_i(0)} < 1 \\ X_i(0) + \Delta_0 X_i & \text{if } \frac{B_i(0)}{X_i(0)} > 1 \end{cases}, \quad i = 1, 2 \tag{E.14}$$

and $G^{(k)}$ and $G_d(\Delta_d X, \mathbf{Z}^d(\Delta_0))$ are given by the equivalent of (E.5). Notice that the set described by equations (E.8) to (E.10) can be re-expressed in a more compact form as follows

$$\boldsymbol{\Delta}_1^{N_1} \in \Gamma_2(\boldsymbol{\Delta}_0^{N_0}) \Longleftrightarrow F(\boldsymbol{\Delta}_0^{N_0}, \boldsymbol{\Delta}_1^{N_1}) \leq 0 \tag{E.15}$$

where $F(\cdot)$ is a continuous function of $(\boldsymbol{\Delta}_0^{N_0}, \boldsymbol{\Delta}_1^{N_1})$. Since $F(\cdot)$ is continuous, for every sequence $\{(\boldsymbol{\Delta}_0^{N_0}(n), \boldsymbol{\Delta}_1^{N_1}(n))\}$ such that $F(\boldsymbol{\Delta}_0^{N_0}(n), \boldsymbol{\Delta}_1^{N_1}(n)) \leq 0$ and $\{(\boldsymbol{\Delta}_0^{N_0}(n), \boldsymbol{\Delta}_1^{N_1}(n))\} \to (\boldsymbol{\Delta}_0^{N_0}(*), \boldsymbol{\Delta}_1^{N_1}(*)))$, $F(\boldsymbol{\Delta}_0^{N_0}(*), \boldsymbol{\Delta}_1^{N_1}(*))) \leq 0$. Given this and (E.15) we conclude that that for every sequence $\boldsymbol{\Delta}_0^{N_0}(n) \to \boldsymbol{\Delta}_0^{N_0}(*)$ and every sequence $\{(\boldsymbol{\Delta}_0^{N_0}(n), \boldsymbol{\Delta}_1^{N_1}(n))\}$ such that $(\boldsymbol{\Delta}_0^{N_0}(n), \boldsymbol{\Delta}_1^{N_1}(n)) \in \Gamma_2(\boldsymbol{\Delta}_0^{N_0}(n))$ for all n, there exists a convergent subsequence of $(\boldsymbol{\Delta}_0^{N_0}(n), \boldsymbol{\Delta}_1^{N_1}(n))$ with limit point $(\boldsymbol{\Delta}_0^{N_0}(*), \boldsymbol{\Delta}_1^{N_1}(*)) \in \Gamma_2(\boldsymbol{\Delta}_0^{N_0}(*))$, i.e., $\Gamma_2(\cdot)$ is upper hemi-continuous.

[4]If $\tilde{\boldsymbol{R}}$ follows a K-nomial process, $\delta_2 \in \mathbb{R}^{12}$.

Moreover, by arguments similar to the ones used in the description of $\Phi_1(\mathbf{Z}(0))$ it is easy to show that for every $\mathbf{\Delta}_0^{N_0}$, $\Gamma_2(\mathbf{\Delta}_0^{N_0})$ is a closed and bounded set. Therefore $\Gamma_2(\cdot)$ is a compact-valued correspondence. We can then describe the feasibility set $\Phi_2(\mathbf{Z}(0))$ as follows

$$\Phi_2(\mathbf{Z}(0)) = \{(\mathbf{\Delta}_0^{N_0}, \mathbf{\Delta}_1^{N_1}) \in \mathbb{R}^{3\times(N_0+N_1)} : \mathbf{\Delta}_0^{N_0} \in \Phi_1(\mathbf{Z}(0)), \mathbf{\Delta}_1^{N_1} \in \Gamma_2(\mathbf{\Delta}_0^{N_0})\} \tag{E.16}$$

i.e., the feasibility set $\Phi_2(\mathbf{Z}(0))$ is the graph of the compact valued upper hemi-continuous correspondence $\Gamma_2(\cdot)$. Since $\Phi_1(\mathbf{Z}(0))$ is compact and $\Gamma_2(\cdot)$ is upper hemi-continuous, it is possible to show (see for example Stokey and Lucas (1989), Exercise 3.15, p.61) that the graph of $\Gamma_2(\cdot)$ is compact and therefore the feasibility set $\Phi_2(\mathbf{Z}(0))$ for the two-period problem is compact.

For the case of a generic T-period problem, let

$$\delta_T = (\mathbf{\Delta}_0^{N_0}, \ldots, \mathbf{\Delta}_{T-1}^{N_{T-1}}) \in \mathbb{R}^{3\times\sum_{t=0}^{T-1} N_t}$$

be a generic investment strategy in the two-period problem. Following the steps above, it is possible to express the feasibility set $\Phi_T(\mathbf{Z}(0))$ as follows

$$\Phi_T(\mathbf{Z}(0)) = \left\{(\delta_T : (\mathbf{\Delta}_1^{N_0}, \ldots \mathbf{\Delta}_{T-2}^{N_{T-2}}) \in \Phi_{T-1}(\mathbf{Z}(0)), \mathbf{\Delta}_{T-1}^{N_{T-1}} \in \Gamma_T(\mathbf{\Delta}_0^{N_0}, \ldots, \mathbf{\Delta}_{T-2}^{N_{T-2}})\right\}$$

where $\Phi_{T-1}(\cdots)$ is a compact set Γ_T is a upper hemi-continuous correspondence. From this we conclude that the feasibility set $\Phi_T(\mathbf{Z}(0))$ is non-empty and compact. ■

Proof of Proposition 3.4.1

Proof: Let us consider $t = T-1$ first. At $t = T-1$ the investor in the two-asset world solves the following problem:

$$V_{T-1}(x_1, x_2) = \max_{\alpha_1, \alpha_2} E\left[\frac{1}{1-\gamma}\Big((x_1+\alpha_1)\tilde{R}_1 + (x_2+\alpha_2)\tilde{R}_2 + (1-x_1-x_2-\alpha_1-\alpha_2)\Big)^{1-\gamma}\right]$$

Let $\alpha^*(x_1, x_2)$ and $\alpha_2^*(x_1, x_2)$ be the solution to this problem. The portfolio holdings in the two asset will hence be

$$\omega(x_1, x_2) = \frac{x_1+\alpha_1}{x_1+\alpha_1+x_2+\alpha_2}, \quad 1-\omega(x_1, x_2) = \frac{x_2+\alpha_2}{x_1+\alpha_1+x_2+\alpha_2} \tag{E.17}$$

Let us now consider an investor whose portfolio allocation at time $T-1$ is $x_m = x_1 + x_2$ and who is allowed to trade a fraction α_m in a mutual fund whose return is defined as follows

$$\tilde{R}_m = \omega\tilde{R}_1 + (1-\omega)\tilde{R}_2.$$

The above investor will solve the following problem

$$V_{T-1}^m(x_m) = \max_{\alpha_m} E\left[\frac{1}{1-\gamma}\left((x_m+\alpha_m)\tilde{R}_m + (1-x_m-\alpha_m)\right)^{1-\gamma}\right]$$

Let $\alpha_m^*(x_m)$ be the solution to this problem.

We now show that $V_{T-1}^m(x_m)$ is equal to $V_{T-1}(x_1, x_2)$ if $\alpha_m^* = \alpha_1^* + \alpha_2^*$. The proof is trivially done by substituting the expression of $\tilde{R_m}$ and x_m.

$$\begin{aligned}
V_{T-1}^m(x_m) &= E\left[\frac{1}{1-\gamma}\left((x_m+\alpha_m^*)\tilde{R}_m + (1-x_m-\alpha_m^*)\right)^{1-\gamma}\right] \\
&= E\left[\frac{1}{1-\gamma}\left((x_1+x_2+\alpha_1^*+\alpha_2^*)\frac{(x_1+\alpha_1^*)\tilde{R}_1+(x_2+\alpha_2^*)\tilde{R}_2}{x_1+x_2+\alpha_1^*+\alpha_2^*}\right.\right. \\
&\quad \left.\left. +(1-x_1-x_2-\alpha_1^*-\alpha_2^*)\right)^{1-\gamma}\right] \\
&= E\left[\frac{1}{1-\gamma}\left((x_1+\alpha_1^*)\tilde{R}_1 + (x_2+\alpha_2^*)\tilde{R}_2 + (1-x_1-x_2-\alpha_1^*-\alpha_2^*)\right)^{1-\gamma}\right] \\
&= V_{T-1}(x_1, x_2)
\end{aligned}$$

Therefore there always exists a mutual fund R_m such that $V_{T-1} = V_{T-1}^m$. Since with CRRA utility function the holding in the risky assets are unchanged through time (i.e. $x_1+\alpha_1^*$ and $x_2+\alpha_2^*$ are the constant Merton's portfolio weights in the two risky assets)[5] the rest of the proof follows by induction on t. ∎

[5]See Merton (1971)

Appendix F

Numerical Implementation for Chapter 3

Gross stock returns are assumed to follow an exogenously specified trinomial process

$$\left(\tilde{R}_{1,t+1}, \tilde{R}_{2,t+1}\right) = \begin{cases} \left(1+(\mu_1-d_1)+\sigma_1\sqrt{\frac{3}{2}}, 1+(\mu_2-d_2)+\sigma_2\rho\sqrt{\frac{3}{2}}+\sigma_2\sqrt{\frac{1-\rho^2}{2}}\right) & \text{with prob. } \frac{1}{3} \\ \left(1+(\mu_1-d_1), 1+(\mu_2-d_2)-2\sigma_2\sqrt{\frac{1-\rho^2}{2}}\right) & \text{with prob. } \frac{1}{3} \\ \left(1+(\mu_1-d_1)-\sigma_1\sqrt{\frac{3}{2}}, 1+(\mu_2-d_2)-\sigma_2\rho\sqrt{\frac{3}{2}}+\sigma_2\sqrt{\frac{1-\rho^2}{2}}\right) & \text{with prob. } \frac{1}{3} \end{cases} \tag{F.1}$$

where μ_i, σ_i, ρ represent the expected return, volatility and correlation coefficient respectively of asset $i = 1, 2$. The above characterization is due to He (1990) who proves that every N-dimensional diffusion price process can be approximated by a sequence of N-variate, $(N+1)$-nomial processes. In particular, He (1990) shows that, for $n \to \infty$, the sequence

$$\left(\tilde{S}_{1,t+1}, \tilde{S}_{2,t+1}\right)^{(n)} = \tag{F.2}$$

$$\begin{cases} \left(S_{1,t}^{(n)} + \frac{(\mu_1-d_1)S_{1,t}^{(n)}}{n} + \sigma_1 S_{1,t}^{(n)}\sqrt{\frac{3}{2n}}, S_{2,t}^{(n)} + \frac{(\mu_2-d_2)S_{2,t}^{(n)}}{n} + \sigma_2\rho S_{2,t}^{(n)}\sqrt{\frac{3}{2n}} + \sigma_2 S_{2,t}^{(n)}\sqrt{\frac{1-\rho^2}{2n}}\right) & \text{w/prob. } \frac{1}{3} \\ \left(S_{1,t}^{(n)} + \frac{(\mu_1-d_1)S_{1,t}^{(n)}}{n}, S_{2,t}^{(n)} + \frac{(\mu_2-d_2)S_{2,t}^{(n)}}{n} - 2\sigma_2 S_{2,t}^{(n)}\sqrt{\frac{1-\rho^2}{2n}}\right) & \text{w/prob. } \frac{1}{3} \\ \left(S_{1,t}^{(n)} + \frac{(\mu_1-d_1)S_{1,t}^{(n)}}{n} - \sigma_1 S_{1,t}^{(n)}\sqrt{\frac{3}{2n}}, S_{2,t}^{(n)} + \frac{(\mu_2-d_2)S_{2,t}^{(n)}}{n} - \sigma_2\rho S_{2,t}^{(n)}\sqrt{\frac{3}{2n}} + \sigma_2 S_{2,t}^{(n)}\sqrt{\frac{1-\rho^2}{2n}}\right) & \text{w/prob. } \frac{1}{3} \end{cases}$$

converges weakly to the following two-dimensional lognormal process

$$\begin{aligned} dS_1 &= (\mu_1-d_1)S_1dt + \sigma_1 S_1 dW_1 \\ dS_2 &= (\mu_2-d_2)S_2dt + \sigma_2\rho S_2 dW_1 + \sigma_2\sqrt{1-\rho^2}S_2 dW_2 \end{aligned}$$

where $S_{i,t}$ represents the stock price at time t, μ_i, d_i, σ_i, ρ represent the instantaneous expected return, dividend yield, volatility and correlation coefficient respectively of asset $i = 1, 2$ and W_1, W_2 are two independent Brownian Motions. This convergence result preserves completeness in the "approximating" markets. The representation in (F.3) is not unique.

Our representation follows by taking the first term of such sequence ($n = 1$) and expressing (F.3) in terms of returns

$$\tilde{R}_{i,t+1} = \frac{S_{i,t+1}}{S_{i,t}}.$$

We discretize the state space $(x_1, x_2, \theta_1, \theta_2)$ in a four-dimensional grid

$$[0, \overline{x}]^2 \times [0, \overline{\theta}]^2$$

and solve the problem by backward recursion. Given the high dimensionality of the problem, in order to gain accuracy, we work with a double four-dimensional grid: an *inner grid*

$$[0, \overline{\overline{x}}]^2 \times [0, \overline{\overline{\theta}}]^2$$

and an *outer grid*

$$[\overline{\overline{x}}, \overline{x}]^2 \times [\overline{\overline{\theta}}, \overline{\theta}]^2, \;\; \overline{\overline{x}} < \overline{x} \text{ and } \overline{\overline{\theta}} < \overline{\theta}.$$

We use a higher precision on the inner grid and a coarser precision on the outer grid. Notice that imposing a bound $\overline{x}$ on the portfolio holdings corresponds to imposing leverage constraints in the problem. Given that we select a high value for $\overline{x}$ and focus on state variables within the inner grid, we minimize the effect of the leverage constraint on the optimal solution.

Return on the Index Fund

In (F.1), if we assume, as in out base case parameters, that the two assets have same mean $\mu = (\mu_1 - d_1) = (\mu_2 - d_2)$, same volatility $\sigma = \sigma_1 = \sigma_2$ and correlation $\rho = 0.5$ then their joint distribution becomes

$$\left(\tilde{R}_{1,t+1}, \tilde{R}_{2,t+1}\right) = \begin{cases} \left(1 + \mu + \sigma\sqrt{\frac{3}{2}}, 1 + \frac{\mu}{n} + \sigma\sqrt{\frac{3}{2}}\right) & \text{with prob. } \frac{1}{3} \\ \left(1 + \mu, 1 + (\mu_2 - d_2) - \sigma\sqrt{\frac{3}{2}}\right) & \text{with prob. } \frac{1}{3} \\ \left(1 + \mu - \sigma\sqrt{\frac{3}{2}}, 1 + \frac{\mu}{n}\right) & \text{with prob. } \frac{1}{3} \end{cases}$$

From the above expression we see that the marginal distribution of the two assets is identical. Therefore the Merton's trades will be identical $\alpha_1^*(x_1, x_2) = \alpha_2^*(x_1, x_2)$, for all (x_1, x_2). The return R_m on the index fund is defined by

$$R_m = \omega R_1 + (1-\omega) R_2, \;\; \omega = \frac{\alpha_1^*(0,0)}{\alpha_1^*(0,0) + \alpha_2^*(0,0)}.$$

When $\rho = 0.5$ $\alpha_1^*(x_1, x_2) = \alpha_2^*(x_1, x_2)$, for all (x_1, x_2) and therefore $\omega = 1/2$, i.e. the index fund is an equally weighted fund.

Given this, the mutual fund consisting of two assets, with $(\mu_1 - d_1) = (\mu_2 - d_2) = \mu$, $\sigma_1 = \sigma_2 = \sigma$ and $\rho = 0.5$ can be implemented as a single asset with return distribution

$$\left(\tilde{R}_{m,t+1}\right) = \begin{cases} 1 + \mu + \sigma\sqrt{\frac{3}{2}} & \text{with prob. } \frac{1}{3} \\ \frac{1}{2}\left(1 + \mu + 1 + (\mu_2 - d_2) - \sigma\sqrt{\frac{3}{2}}\right) & \text{with prob. } \frac{2}{3} \end{cases}$$

Appendix G

Tables and Figures for Chapter 3

Table XII. Benchmark Case Parameter Values for the Numerical Solution of the Two-Asset Case. The investment horizon is T. The parameters μ_i, d_i, and σ_i give the cum-dividend mean, dividend yield, and standard deviation of asset i returns for $i = 1, 2$. The correlation coefficient is ρ. The tax rate on capital gains and dividends are τ_g and τ_d respectively. The risk aversion parameter is γ, and the subjective discount factor is β.

Parameter	Value
T	5
μ_1	0.10
μ_2	0.10
d_1	0.00
d_2	0.00
σ_1	0.30
σ_2	0.30
ρ	0.50
τ_g	0.25
τ_d	0.25
γ	5.00
β	0.96

Table XIII. Value of Tax-Deferral Option. The table reports the value of the tax-timing option for different levels of the volatility of the two risky assets and for different level of risk aversion. The value is computed as the extra percentage wealth necessary to make an individual paying taxes on accrual indifferent to holding the tax-on-realization portfolio. Fro every case, the option value is computed at the point in which the investor holds the Merton's portfolio and the assets do not embed any gain or loss ($\theta_1 = 1, \theta_2 = 1$). Other parameters are as in Table XII.

$\sigma_1 = \sigma_2$	$\gamma = 3.0$	$\gamma = 5.0$	$\gamma = 7.0$
0.25	7.50%	4.11%	2.78%
0.30	5.61%	3.05%	2.07%
0.35	4.37%	2.39%	1.61%

Table XIV. Value of Tax-Deferral Option. The table reports the value of the tax-timing option for different levels of correlation between the two risky assets. The value is computed as the extra percentage wealth necessary to make an individual paying taxes on accrual indifferent to holding the tax-on-realization portfolio. Asset holdings are set such that 20% of wealth is held in each asset (i.e. the Merton portfolio, $x_1 = .2, x_2 = .2$) and the basis to price ratio is $\theta_1 = 1, \theta_2 = 1$. Other parameters are as in Table XII.

ρ	Option Value
-0.50	12.27%
0.00	5.11%
0.50	3.05%
0.90	1.98%

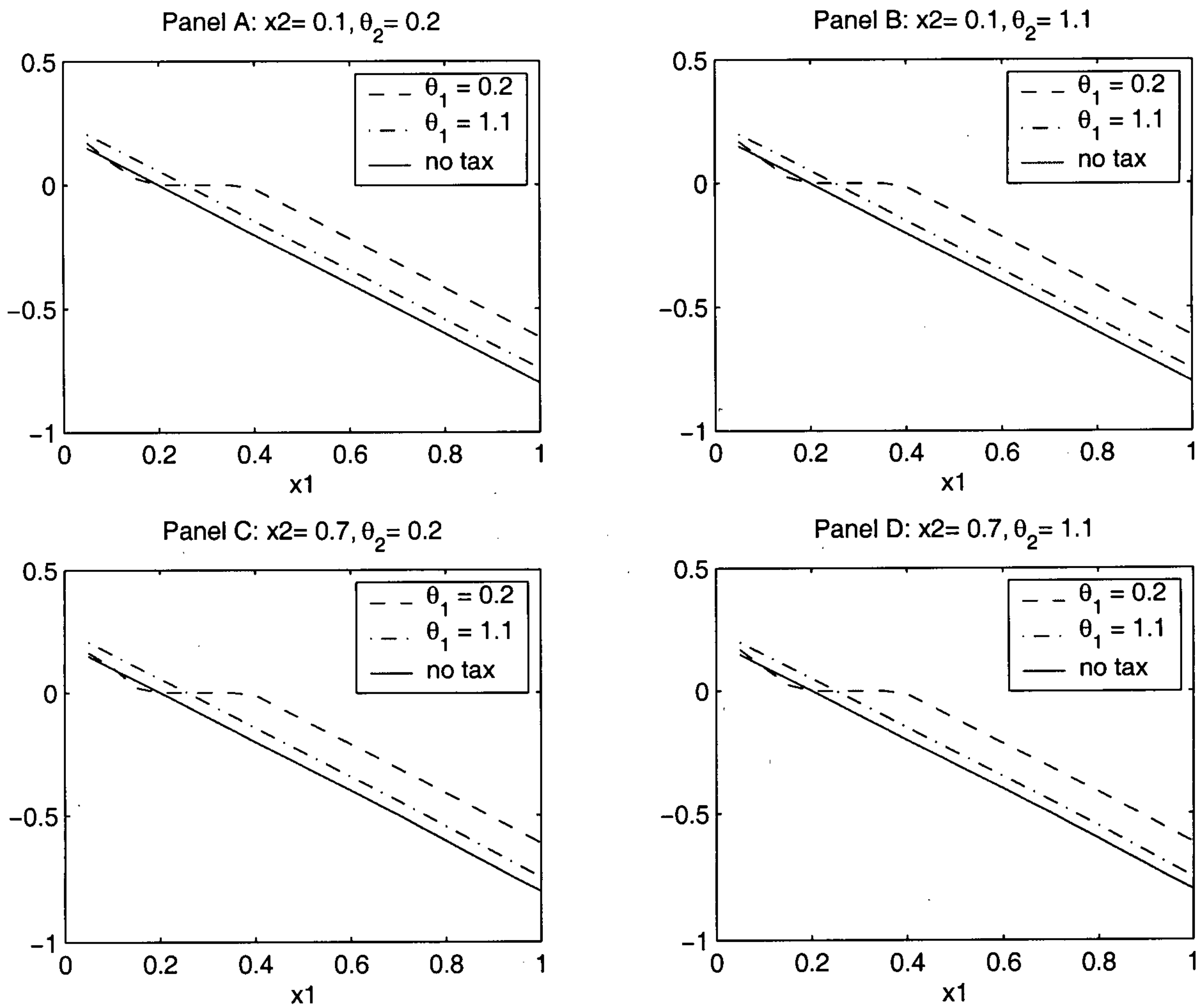

Figure G.1. Trades in Stock 1. This figure shows trades in asset 1 as a function of holdings in asset 1 for a variety of values of the other state variables. Each panel reports the trades for two different levels of the basis to price ratio in asset 1 ($\theta_1 = 0.2, 1.1$) as well as the trades in the absence of taxes. Panel A reports a situation in which asset 2 has a low holding and an embedded gain. Panel B reports a situation in which asset 2 has a low holding and an embedded loss. Panel C reports a situation in which asset 2 has a high holding and an embedded gain. Panel D reports a situation in which asset 2 has a high holding and an embedded loss. Parameter values are as specified in Table XII.

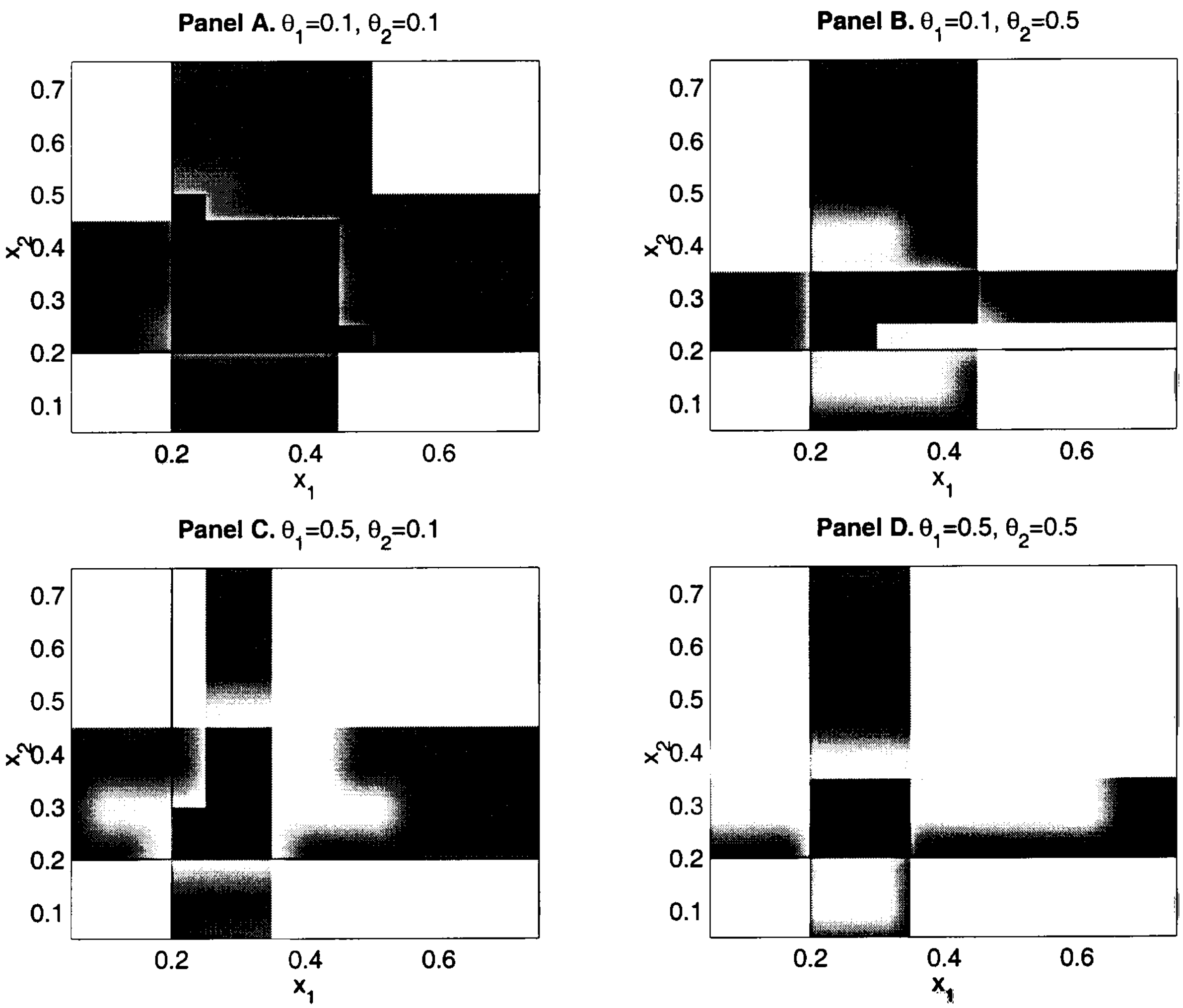

Figure G.2. Regions of No-Trade. This Figure shows the investment strategy at time $t = 0$ in the two asset problem for a variety of different basis to price ratios. Three different types of actions are detailed as a function of the initial portfolio positions in each asset. In the black region, the investor chooses not to trade in either asset. In the grey region, one asset is traded (bought or sold), while holdings of the other asset remain constant. The white region shows values of the initial holdings where the investor chooses to trade in both assets. The black cross indicates the investment by the non-taxed (Merton) investor.

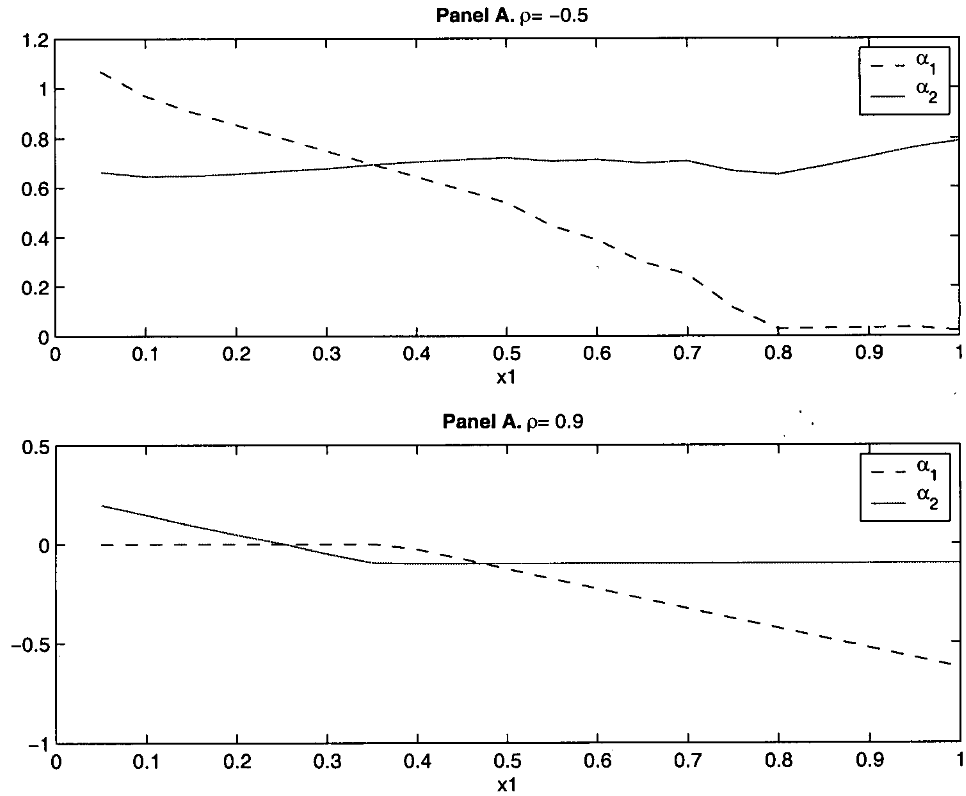

Figure G.3. Cross-effects of holdings on trades. This figure shows the trades in asset 1 (α_1) and asset 2 (α_2) as a function of the holding of asset 1 (x_1). For both panels, the holdings of asset 2 are fixed at $x_2 = 0.1$, the basis to price ratio for asset 1 is $\theta_1 = 0.2$ and the basis to price ratio for asset 2 is $\theta_2 = 1.0$. Panel A refers to the case where the two assets are negatively correlated ($\rho = -0.5$), while Panel B refers to the case where the two assets are nearly perfect substitutes ($\rho = 0.9$). The other parameter values are as specified in Table XII.

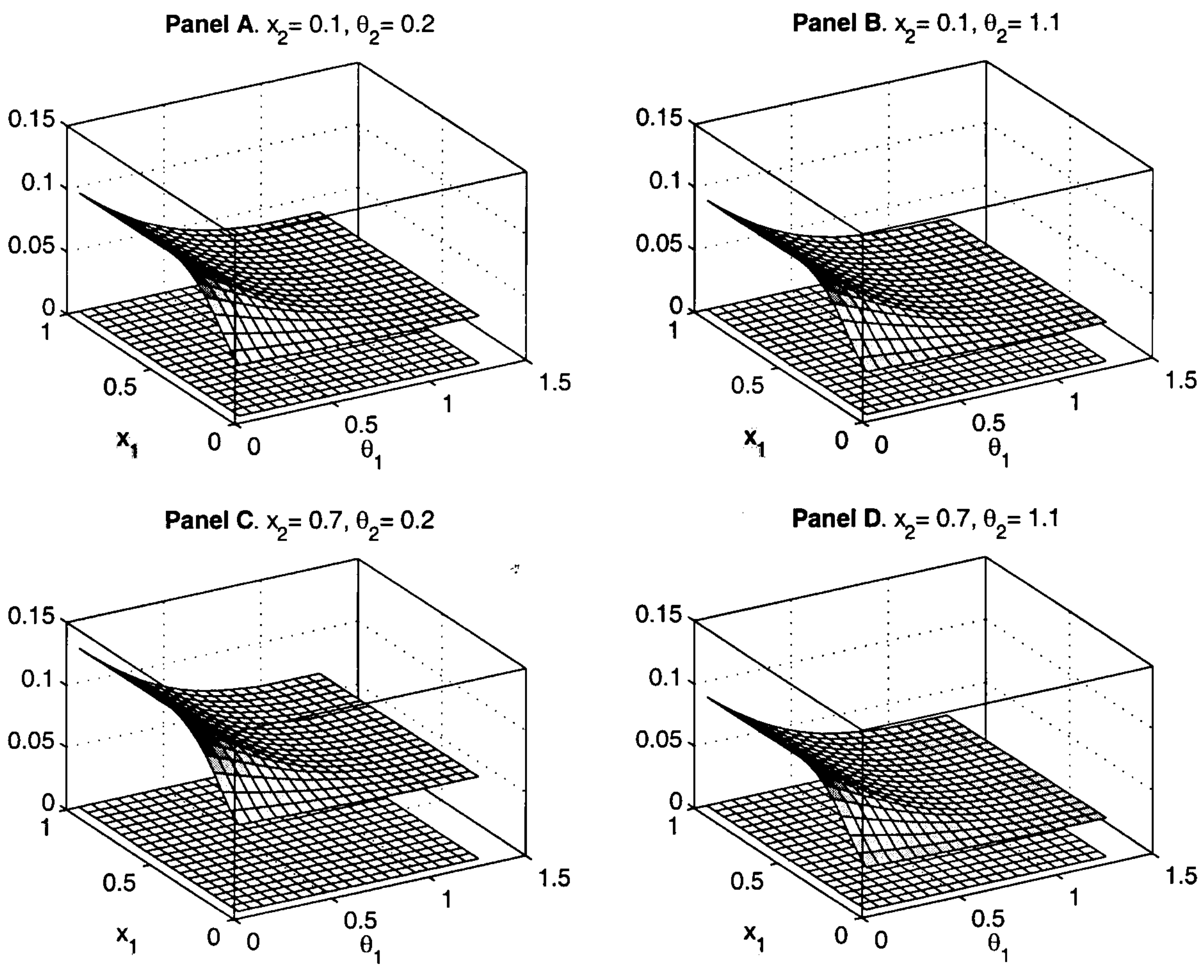

Figure G.4. Value of Tax-Timing Options. The value of the tax-timing option is computed as the extra percentage wealth necessary to make the no-tax investor indifferent to holding the taxable portfolio. The figure plots values as a function of the basis and initial holding in asset 1. Every panel plots the option value as a function of holdings in asset 1 (x_1) and its relative basis to price ratio (θ_1). Panel A shows the case with low holdings and a capital gain in asset 2. Panel B shows the case with low holdings and a capital loss in asset 2. Panel C shows the case with high holdings and a capital gain in asset 2. Panel D shows the case with high holdings and a capital loss in asset 2. Parameter values are as specified in Table XII.

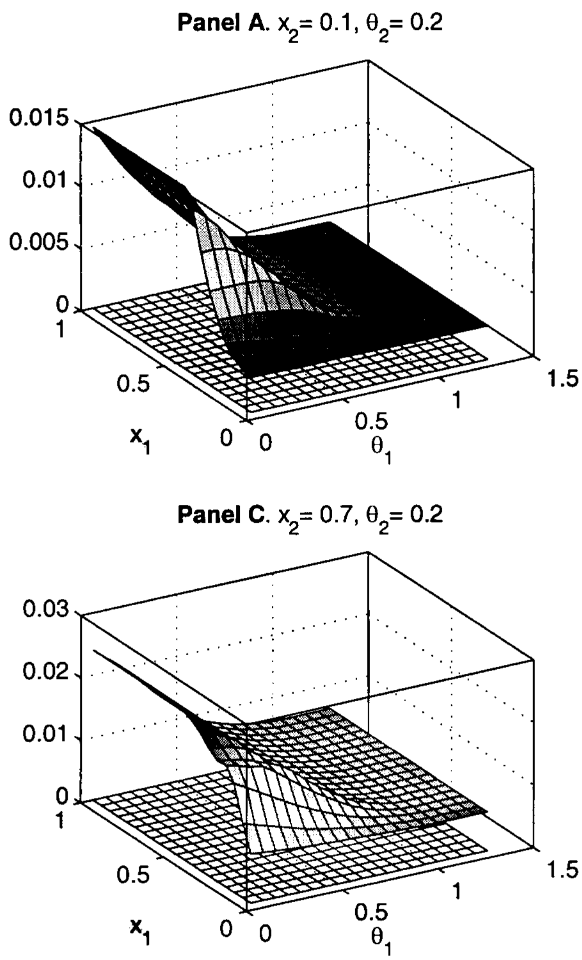

Figure G.5. Value of multi-asset flexibility. This Figure shows the certainty equivalent cost of investing in a mutual fund which which passes all gains and losses to the shareholder. Certainty equivalent costs are defined as the percentage additional wealth necessary to make the mutual fund investor indifferent to holding the two taxable assets individually. Positive values indicate states where the two asset case is preferred. Every panel plots the certainty equivalent costs as a function of holdings in asset 1 (x_1) and its relative basis to price ratio (θ_1). Panel A shows the case with low holdings and a capital gain in asset 2. Panel B shows the case with low holdings and a capital loss in asset 2. Panel C shows the case with high holdings and a capital gain in asset 2. Panel D shows the case with high holdings and a capital loss in asset 2. Parameter values are as specified in Table XII.

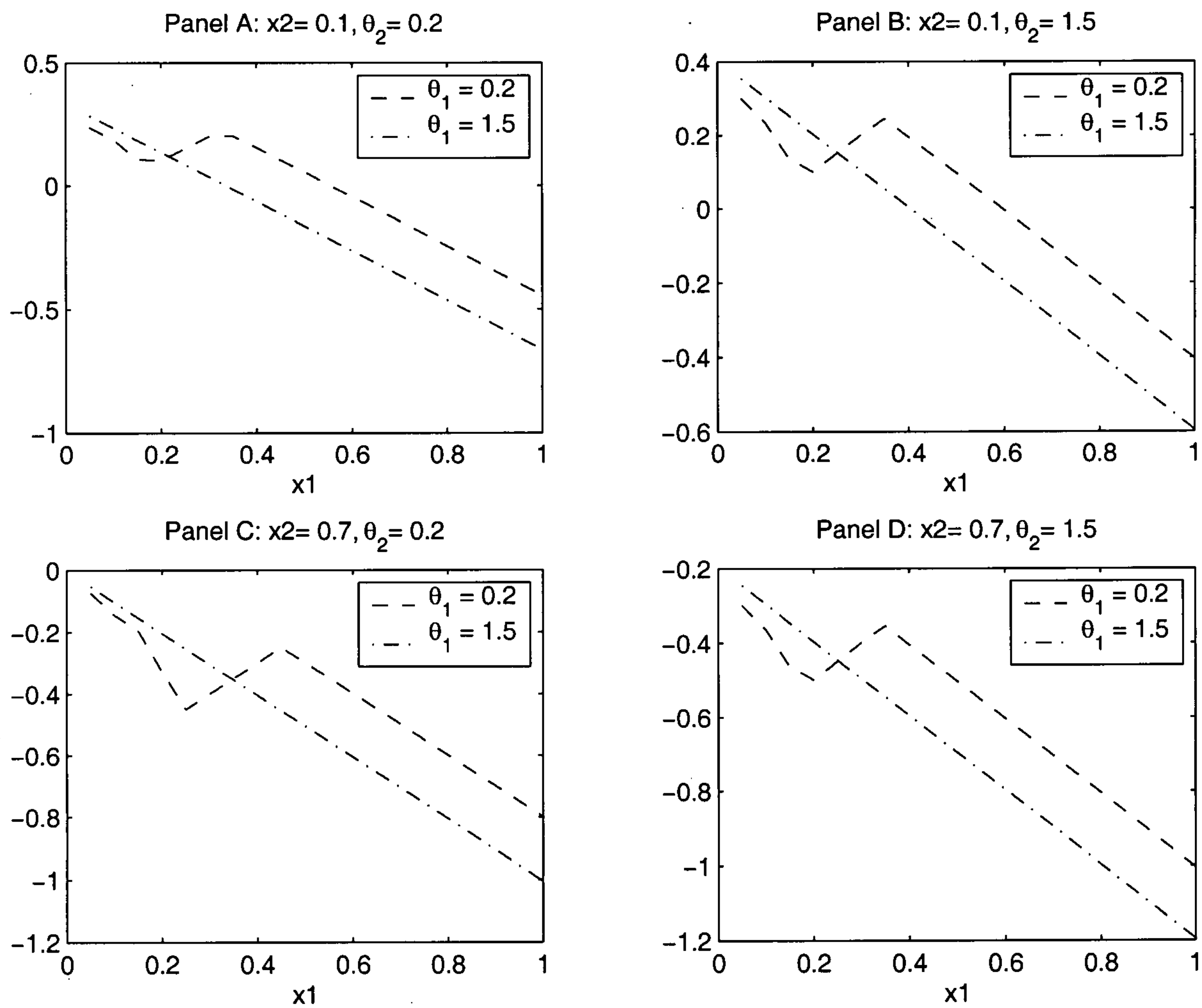

Figure G.6. Trades in the Mutual Fund Units. This figure shows trades in the mutual fund (α_m) as a function of holdings in asset 1 for a variety of values of the other state variables. Each panel reports the trades for two different levels of the basis to price ratio in asset 1 ($\theta_1 = 0.2, 1.5$). Panel A reports a situation in which asset 2 has a low holding and has an embedded gain. Panel B reports a situation in which asset 2 has a low holding and an embedded loss. Panel C reports a situation in which asset 2 has a high holding and an embedded gain. Panel D reports a situation in which asset 2 has a high holding and an embedded loss. Parameter values are as specified in Table XII.

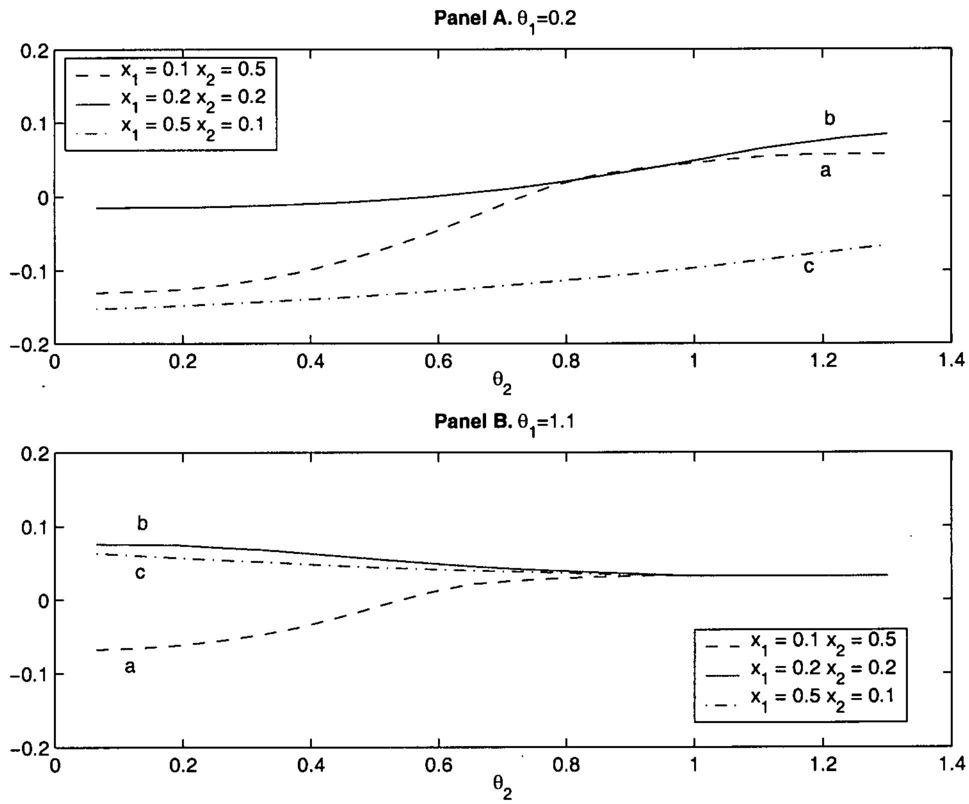

Figure G.7. Value of Index Fund Trading. This Figure shows the certainty equivalent costs of investing in an index fund which which has tax treatment equivalent to a composite asset in contrast to investing in the two assets separately. Certainty equivalent costs are defined as the amount of additional wealth necessary to make the index fund investor indifferent to holding the two taxable assets individually. The certainty equivalent costs are normalized by initial wealth. Positive values indicate states where the two asset case is preferred. Each panel reports the certainty equivalent cost for three different portfolio allocations as the basis to price ratio in asset 2 (θ_2) varies. Line a refers to a portfolio overbalanced in asset 2. Line b refers to a portfolio equally weighted in the two assets and line c refers to a portfolio overbalanced in asset 1. Panel A shows the case of an embedded gain in asset 1 and Panel B shows the case of an embedded loss in asset 1. Parameter values are as specified in Table XII.

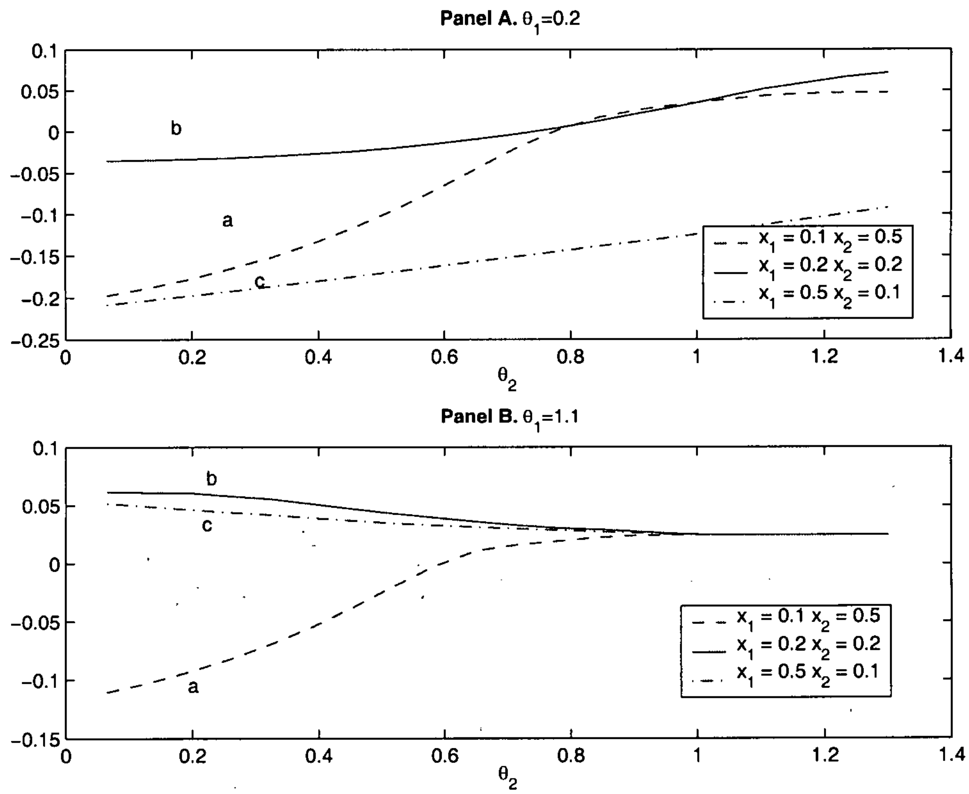

Figure G.8. Tax Costs of Mutual Fund Turnover. This Figure shows the certainty equivalent cost of investing in an index fund which which has tax treatment equivalent to a composite asset in contrast to investing in a mutual fund which passes all realized gains and losses. Certainty equivalent costs are defined as the amount of additional wealth necessary to make the index fund investor indifferent to holding the mutual fund. The certainty equivalent cost are normalized by initial wealth. Positive values indicate states where the mutual fund case is preferred. Each panel reports the certainty equivalent cost for three different portfolio allocations as the basis to price ratio in asset 2 (θ_2) varies. Line a refers to a portfolio overbalanced in asset 2. Line b refers to a portfolio equally weighted in the two assets and line c refers to a portfolio overbalanced in asset 1. Panel A shows the case of an embedded gain in asset 1 and Panel B shows the case of an embedded loss in asset 1. Parameter values are as specified in Table XII.

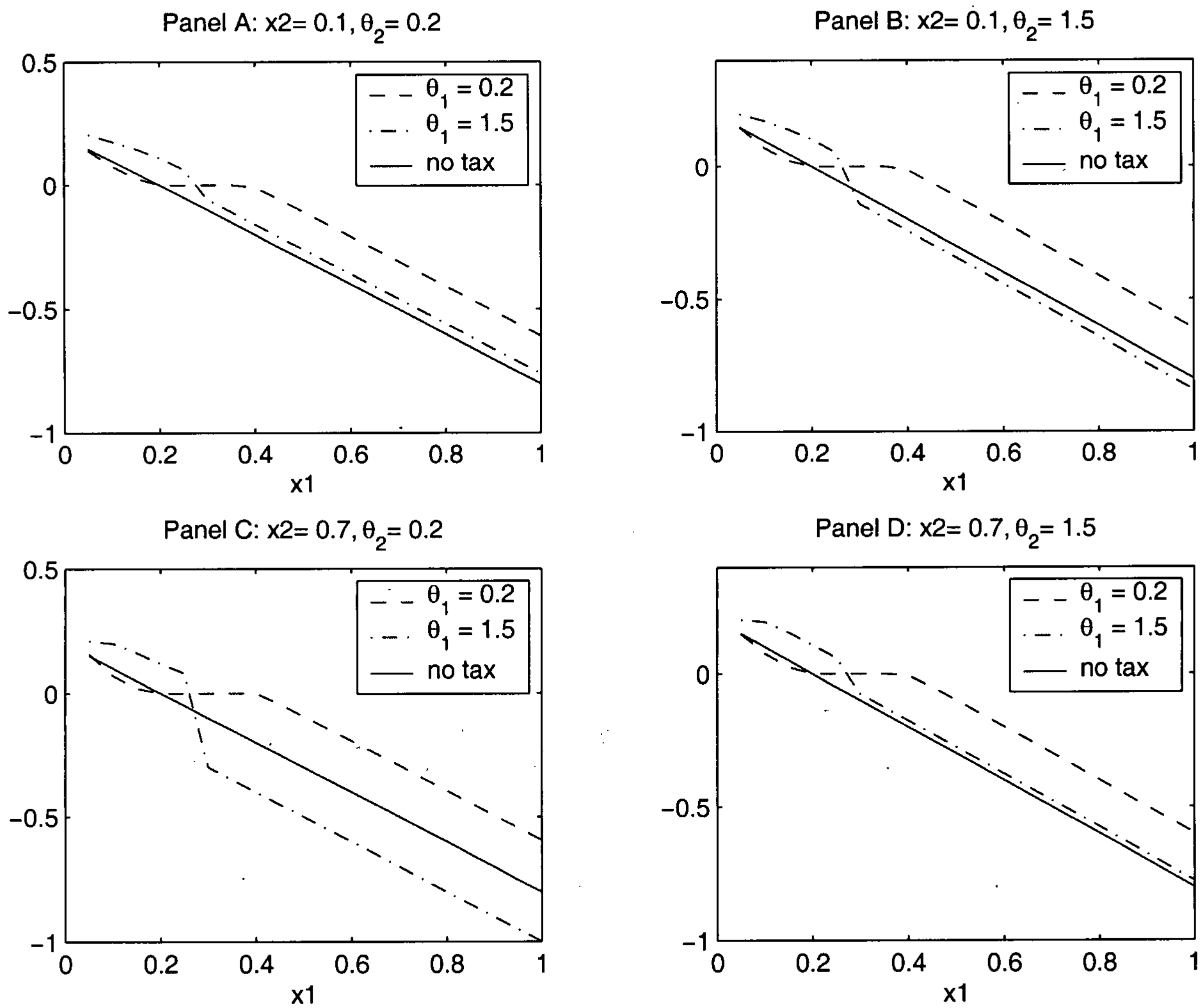

Figure G.9. Trades in Stock 1 Under the Wash Sales Constraint. This figure shows trades in asset 1 as a function of holdings in asset 1 when wash sales are not allowed and for a variety of values of the other state variables. Each panel reports the trades for two different levels of the basis to price ratio in asset 1 ($\theta_1 = 0.2, 1.5$) as well as the trades in the absence of taxes. Panel A reports a situation in which asset 2 has a low holding and an embedded gain. Panel B reports a situation in which asset 2 has a low holding and an embedded loss. Panel C reports a situation in which asset 2 has a high holding and an embedded gain. Panel D reports a situation in which asset 2 has a high holding and an embedded loss. Parameter values are as specified in Table XII.

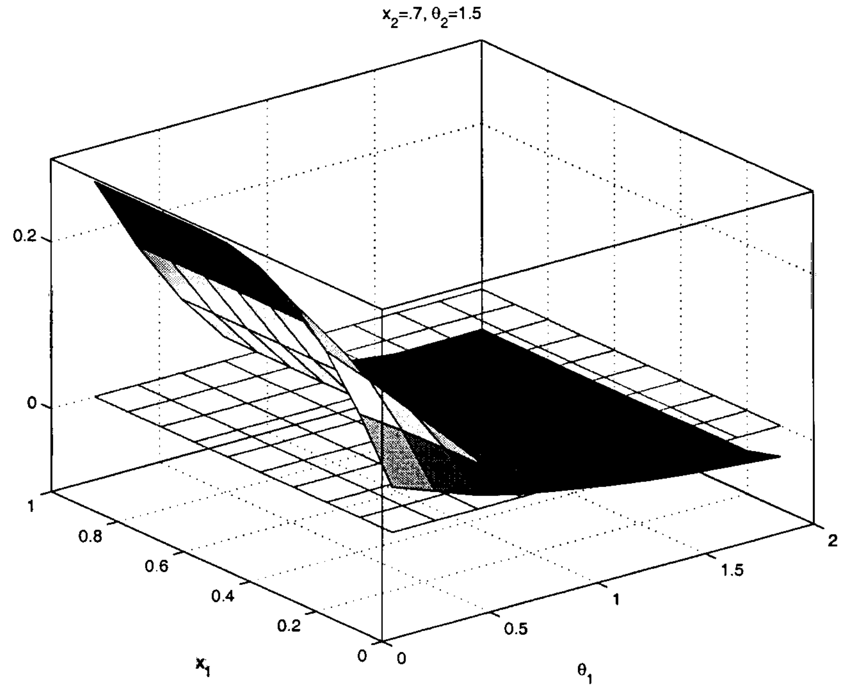

Figure G.10. Value of Tax-timing option when wash sales are not allowed. This Figure shows the certainty equivalent cost of investing in two risky assets in the presence of capital gains taxes compared to the no-tax case. Wash sales are not allowed. Certainty equivalent costs are defined as the amount of additional wealth necessary to make the non-taxable investor indifferent to holding assets subject to taxation. Positive values indicate states where the two asset case is preferred. The figure is drawn for the case of an initial holding of asset 2 equal to $x_2 = 0.7$ and for a basis to price ratio $\theta_2 = 1.5$. Parameter values are as specified in Table XII with the exception of the horizon which is set to T=2.

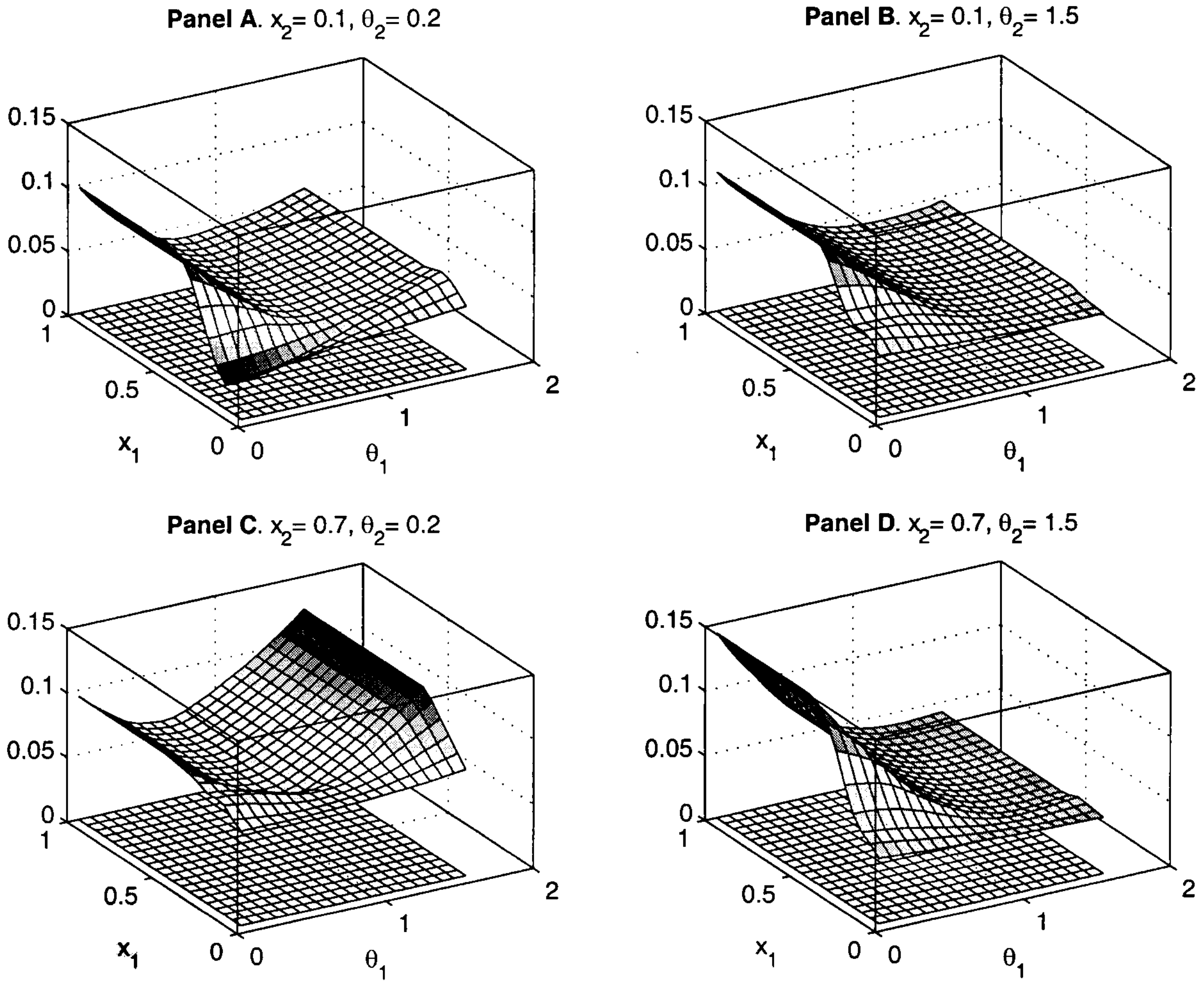

Figure G.11. Value of multi-asset flexibility when wash sales are not allowed. This Figure shows the certainty equivalent cost of investing in a mutual fund which which passes all gains and losses to the shareholder. Wash sales are not allowed. Certainty equivalent costs are defined as the amount of additional wealth necessary to make the mutual fund investor indifferent to holding the two taxable assets individually. Positive values indicate states where the two asset case is preferred.. Every panel plots the certainty equivalent costs as a function of holdings in asset 1 (x_1) and its relative basis to price ratio (θ_1). Panel A shows the case with low holdings and a capital gain in asset 2. Panel B shows the case with low holdings and a capital loss in asset 2. Panel C shows the case with high holdings and a capital gain in asset 2. Panel D shows the case with high holdings and a capital loss in asset 2. Parameter values are as specified in Table XII.