Essays in Asset Pricing and Portfolio Choice

by

Oleg Shibanov

A thesis submitted for the degree of

Doctor of Philosophy

Department of Finance

London Business School

May 2011

Declaration

I, Oleg Shibanov, hereby declare that the work in this dissertation was carried out in accordance with the Regulations of London Business School. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. Any views expressed in the dissertation are those of the author and in no way represent those of London Business School. The dissertation has not been presented to any other University for examination either in the United Kingdom or overseas.

Signature:

Date:

Acknowledgements

I am profoundly indebted to my supervisor Francisco Gomes. This thesis would not have been possible without his unlimited support, encouragement and advice. I would like to especially thank my transfer committee member Anna Pavlova, who has supported me through my PhD studies and has given me valuable feedbacks on my research. I would like to thank Suleyman Basak, who was a PhD coordinator during the first and second year at PhD Programme, for his valuable guidance, advice on my academic goals and important feedback on my research. I am obliged to Ksenia Panidi and Federica Pievani for their support and beneficial comments. I am also grateful to Andrea Buffa, Georgy Chabakauri, Joao Cocco, Ian Cooper, Elroy Dimson, Vito Gala, Joao Gomes, Jungsuk Han, Igor Makarov, Albert Marcet, Helene Rey, Andrew Scott and Henry Servaes, for helpful discussions.

Finally, I would like to give special thanks to my parents for their constant and endless support.

Abstract

This thesis is devoted to studying portfolio choice and cross-section of returns in three different frameworks.

In the first chapter I study the impact of a global cap-and-trade system for greenhouse gas emissions on financial markets. Firms face shrinking caps on their emissions and can reduce these emissions by installing a cleaner technology. Two main results are obtained. First, producers that have lower emissions at the start-up date have consistently higher expected returns on equity. Second, the higher the fraction of free permits and the reduction in the emissions after installment of the cleaner technology, the lower the spread between returns. I obtain preliminary empirical results that support the implications of the model.

In the second chapter I study the relation between fees and performance in the U.S. mutual fund industry. I show that not only risk-adjusted before-fee return (alpha) but also its volatility (sigma) have significant impact on fees and fund flows. Three novel results are obtained. First, the level of fees is positively related to both alpha and sigma. Second, the change in fees is positively related to sigma before 2000 yet negatively after 2000, and is negatively related to alpha. The latter result may seem counterintuitive, and I further show that increase in fees is followed by improved performance. Finally, fund flows depend positively on sigma. I rationalize these results in a simple model.

In the third chapter I study optimal labor, consumption, training and portfolio decisions in a life-cycle model with human capital and wealth accumulation. The agent can increase his future earnings by augmenting the human capital through training and learning-by-doing. It is shown that the levels of wage income and the shape of wealth can be matched to the data, while the share of risky asset in the portfolio exhibits an inverse U-shaped form.

Contents

List of Figures

List of Tables

Introduction

In this thesis I study optimal individual choice, including portfolio choice, in three distinct frameworks. In each of three chapters I consider different economic stories, yet my main aim is to find the impact of fundamental or idiosyncratic risks on portfolio choice. In the first chapter I study riskiness of firms under a cap-and-trade system. In the second chapter I show the influence of higher moments of risk-adjusted returns on mutual funds decisions. In the third chapter I derive optimal portfolio allocations in the case of endogenous human capital accumulation.

In the first chapter I study the impact of a global cap-and-trade system for greenhouse gas emissions on financial markets. The analysis is conducted using a general equilibrium production economy with two countries. One of the countries (developed country) introduces the system from the very beginning while the other country (developing country) can join it later. Firms in the first country are divided into two types with either high or low emission levels. Therefore, this setup allows me to study both cross-country and within-country differences. Producers face shrinking caps on their emissions and can reduce these emissions by installing a cleaner technology. Three main results are obtained in this setting. First, producers that have lower emissions at the start-up date have consistently higher expected returns on equity. Second, this return spread depends on the fraction of emission permits given for free, and on the efficiency of the cleaner technology. The higher the fraction of free permits and the reduction in the emissions after installment of the cleaner technology, the lower the spread. Further, I show that if the second country joins the cap-and-trade agreement at the start-up date, and if it is given a high enough number of permits, both countries are better off in terms of consumption streams. But they might be worse off in terms of emissions after taking into account the long-lasting impact of the greenhouse gases. So, I derive the level of permits that makes the second country break-even in terms of consumption stream and show that this level leads to higher long-term emissions than in the case of the first country with caps only. The results are stable to a range of parameters. I obtain preliminary empirical results that support the implications of the model.

Second chapter is devoted to the study of the relation between fees and performance in the U.S. mutual fund industry. I consider the relation between fees, fund flows, risk-adjusted performance ("alpha") and its volatility ("sigma"). I find several novel empirical facts. First, the level of fees is positively related both to alpha and sigma. Second, fund flows increase both in sigma and alpha. I also show that when funds alter their fees, this change impacts fund flows: increase in fees leads to lower fund flows for the next one to three months, and reduction in fees lead to higher fund flows. Third, I show that the change in fees depends positively on past sigma before

2000 but depends negatively on past sigma starting from 2000. Moreover, the change in fees depends negatively on past alpha. The latter result may look counterintuitive as it means that the funds with lower past performance increase fees. To investigate it further, I show that increase in fees is followed by a better performance with respect to the last year. Yet this result is a time-series one: in the cross-section, funds that increase fees do not show improved returns in comparison to their competitors. Finally, I rationalize the results in a simple model.

In the third chapter I study optimal labor, consumption, training and portfolio decisions in a life-cycle model with human capital and wealth accumulation. The agent can increase his future earnings by augmenting the human capital through training and learning-by-doing. The framework allows for retirement capital and i.i.d. process for employment and assumes two types of agents: constrained, who cannot invest in risky stock, and unconstrained. It is shown that the levels of wage income and the shape of wealth can be matched to the data, while the share of risky asset in the portfolio exhibits an inverse U-shaped form. Two special cases in which the agents either invest half of the wealth into risky stock or work for half of available time are studied.

All remaining errors are mine.

1. Financial Markets and Cap-and-Trade System in a General Equilibrium Model

1.1. Introduction

"Global warming" and a "green economy" have become hot topics in recent years. While there is a huge debate on the ability of economic models to forecast the future of climate change (see, e.g., Ackerman et al. (2009); Bansal and Ochoa (2009a, 2009b); Newbold et al. (2009); Rabl and van de Zwaan (2009); Weitzman (2009a, b, c)), there is little doubt that over time both U.S. and developing countries will join a protocol to reduce their emissions. Although developing countries, like China and India, oppose this type of plan due to the harm it might do to their growth prospects, it seems highly likely that, at some point, they will join a Kyoto-type agreement. As Carey (2010) puts it:

OECD (2009) analysis ... suggests that, in the absence of participation by the United States and any other large emitter, it would be difficult to form a coalition of countries and regions capable of achieving the ...[emission] target by 2050 through a single (coalition-wide) feasible carbon price without mitigation costs becoming very high and would be physically impossible if neither the United States nor China participated (other countries would have to have negative emissions).

Under a cap-and-trade system every participating country agrees to reduce the emission of greenhouse gases (GHG), mostly CO_2, during a given period of time that is usually two to four years. Under this system, firms are given a limited amount of free permits to emit and if they exceed this limit they have to buy additional allowances to cover their emissions either from other firms involved in the system or from the government. Other ways to reduce emissions include offsets (investing in the reduction of emission outside the country) or the installment of cleaner technology, for example, cleaner fuel. Firms that do not meet the target must pay a fee that can be as high as \$100 per ton of carbon (while the market price of permits is around

\$20 per ton of carbon). The ultimate goal of the cap-and-trade system is not only to reduce emissions but also to stabilize the concentration of CO_2 in the atmosphere at acceptable levels. The cap-and-trade system is the most often used mechanism to implement CO_2 emissions reduction: it is proposed in the U.S. and constantly used in Europe.

While most of the research focuses on the economic implications of the introduction of caps on greenhouse gases, little deals with the financial implications, and none addresses financial markets and technological improvements simultaneously. The goal of this chapter is twofold: to link financial markets with a cap-and-trade system and to introduce technological choice for the firms. To my knowledge, I am the first to try to develop a theoretical model that addresses these two questions in a unified framework.

There exist a link between climate change and asset allocation. As recent Mercer (2011) report shows, climate change would have very different impact on different asset classes and geographic regions. Thus, strategic investors could take it into account, and their portfolio composition should reflect additional risk arising from uncertain climate scenarios. My study addresses some of the fundamental risks a firm participating in a cap-and-trade system faces.

There is a stream of literature that estimates the economic effects of a cap-and-trade system (or, in general, the carbon emission permit-trading system) on economic variables in a computational dynamic general equilibrium setting. Papers of this type include McKibbin et al. (1999), Bosetti et al. (2006, 2008) and OECD (2009). These papers investigate the implications of delays and uncertainties in the participation of developing countries in a climate agreement. Another stream of literature discusses the difference between alternative mechanisms of emission cap introduction and their relation to technology diffusion, e.g., Baker et al. (2008) and Coria (2009). Based on the latter, I concentrate on permit markets and ignore other types of emission mechanisms like a carbon tax.

Bansal and Ochoa (2009a, 2009b) introduce temperature risks with long-term impacts on the prices and returns of financial assets. They incorporate small-probability, high-damage events into the Bansal and Yaron (2004) framework and show that the model is able to match consumption dynamics and moments of financial variables. The authors also estimate expected loss from disasters in terms of consumption. Weitzman (2009a, b, c) shows that if one attempts to quantify the damage born by small-probability disaster events, one observes that the distribution of the probability of these events might have fat right tails and it might be beneficial to start investing in carbon emission reduction immediately.

This chapter also builds on asset pricing in a production economy framework (e.g., Jermann (1998); Boldrin et al. (2001); Balvers et al. (2007); Gomes and Michaelides (2008); Belo (2010); Campanalea et al. (2010)) and has a relation with international finance literature because I study a two-country model (e.g., Ravn et al. (2006); Pavlova and Rigobon (2007); Heyerdahl-Larsen (2010); Grishchenko (2010)).

Further, there are several recent empirical papers on the implications for financial markets and economic performance of a cap-and-trade system. Oberndorfer (2009) and Veith et al. (2009) study the relation between returns on equity for an electricity/power sector and the price of emission permits in the EU-ETS market. They show that there exists a positive relation between the two and that the relation is country and time dependent. This relation might provide some evidence that investors expect the increase in electricity prices to compensate for the higher costs of production. However, the EU ETS market lacks stationarity during this period (see Figure 1.1) and, thus, these studies might lead to spurious regressions. Bushnell et al. (2009) provide a case-study of the April 2006 decrease in value for almost all power plants following an unexpected crush in the price of permits. They show that the cleanest firms in the industries, which were net short-long in permits, experienced the deepest decline; while the firms with the most allowances in the industries, which were net long in permits, were most harmed. Martin et al. (2009) estimate the impact of two types of regulation for carbon emissions in the UK. They observe no significant difference for economic growth between firms taxed for their greenhouse gas emissions and firms that participated in a cap-and-trade system. In a study of NO_x gas regulation in the U.S. during the late 90s to 2004, Fowlie (2010) shows that asymmetric regulation leads to a worse outcome than a more uniform one would: a larger share of pollution was made in states with the worst air quality. This result raises the question of the optimal regulation for the polluters.

In this chapter, I consider a two-country model with four economic sectors and the installment of clean technology. I assume that one of these four sectors produces fuel (oil or coal) that is supplied to the second sector, power plants. Electricity produced in the latter is then used in the production of both consumption goods and investment goods. I start by imposing the cap on the emission of power plants, which seems to be the most plausible scenario under current U.S. regulation. It is assumed that the developed country introduces the cap from the very beginning, and the developing country joins the system later. Two components allow me to produce a plausible risk premium. First, consumers in the model have (deep) habit formation as in Ravn et al. (2006). Second, there exist frictions in the capital market that make the supply of capital not perfectly elastic; and also in the labor market that prevent firms from rebalancing their labor demand immediately after the realization of current productivity shocks (Jermann (1998); Boldrin et al. (2001)). These features do not allow the agent to smooth the labor supply and help to produce a high risk premium. The model replicates several of the main asset-pricing and macroeconomic statistics of the data such as the risk premium, risk-free rate, labor supply, and consumption growth volatility. In the case with the introduction of a cap-and-trade system in the developed country only, I show that the "cleaner" firms (firms with the lowest initial emissions) have higher expected returns. This cross-sectional result is the implication of the more pro-cyclical cash flows of a cleaner firm. During a boom (higher productivity) it has to buy less permits (or can sell more permits) than a "dirtier" firm and, hence, has higher cash flows. In a recession, both cleaner and dirtier firms decrease their production and spend less on permits but this decrease is lower than the increase in a boom. The higher correlation of cash flows with the stochastic discount factor means higher risk and thus higher expected returns.

In addition, I study the dependency of the return spread between cleaner and dirtier firms on the fraction of emission permits given for free (the rest is auctioned), and on the efficiency of the cleaner technology. I show that two results hold. First, the higher the fraction of free permits, the lower the spread. Second. the lower the reduction in the emissions after installment of the cleaner technology, the higher the spread.

When the developing country joins the system, the cross-sectional results are preserved but I have additional considerations. First, what should be the initial allowances given to the developing country? If the number of permits is low, there is no incentive for it to join the system because it might only reduce its growth. On the other hand, if the number of permits is high then there is no reduction in emissions initially. I derive the level of permits that makes the developing country indifferent to joining or not joining and show that even this level leads to a higher cumulative emissions than the baseline case.

I attempt to find empirical support for my model and provide some preliminary results. It seems that the results are in line with the predictions: net buyers of permits have lower *realized* returns than the net sellers, and market betas for the former *decrease more* after the introduction of a cap-and-trade system. Moreover, correlation of realized returns with consumption growth is *higher* for the net sellers of permits.

The structure of the chapter is as follows. In Section 1.2, I describe the model and discuss the solution strategy. Section 1.3 contains the parameters of calibration of the model, and section 1.4 contains a discussion of the results. Preliminary empirical results are reported in section 1.5. Section 1.6 concludes. Appendix 1.7 provides solution of the model, and Appendix 1.8 contains the proof for the optimal installment of the cleaner technology.

1.2. The Model

1.2.1. Firms

I consider a production economy with two countries and representative consumers in each country. Country A represents developed countries, and country B represents developing countries. Each country has four sectors in its economy. The first (consumption sector) and second (investment sector) consist of one representative firm each that produces consumption and investment goods, respectively. There are two consumption goods produced; one in country A (good A) and the other in country B (good B). These two sectors play the key role in the creation of the risk premium. Following Boldrin et al. (2001), I assume that the choice of capital and labor in both sectors for the next period occurs before the realization of the current shock to productivity. Because of that, capital supply is not perfectly elastic and thus the volatility of the return on capital is high. Similarly, labor frictions prevent smoothing in consumption and help to produce a high-risk premium. The third element of the risk premium creation is the deep habits in consumption.

The third sector (electricity sector) of country A consists of two firms that differ in their ability to capture CO_2 and other GHG, but there is only one firm in the respective sector of country B. Firms in the electricity sector produce electricity that is used in the consumption and investment sectors. The fourth sector (fuel) provides fuel to the third sector. I assume that the investment, electricity, and fuel sectors are all local and there is no global trading of the goods produced in these sectors.

The economy has an infinite horizon and quarterly periods, and I assume that all the dynamics in emission caps and technology installment take place in the first 45-year period. Thereafter, the economy is the same from the emission caps perspective: there is no change in emission functions, no change in caps, and cleaner technology installment is free. This assumption simplifies my analysis while leaving me enough freedom to study the effects of interest.

Consumption-good and investment-good firms produce via a production function of the form

$$G^i(a_t, l_t) = a_t^{\psi}(Z_t^i l_t)^{(1-\alpha)(1-\psi)} k_t^{\alpha(1-\psi)}, \quad i = \{A, B\}, t = 1, 2, 3..., \tag{1}$$

where l_t is the labor supply, k_t is the capital stock, and a_t is the amount of electricity used by the firm i. Z_t^i is the productivity of the firms in country i and follows the mix of a zero-mean AR(1) process and deterministic growth:

$$Z_t^i = Z_{t-1}^i e^{\mu_t^i + \zeta_t^i}, i = \{A, B\}, \; \zeta_t^i = \rho_0 \zeta_{t-1}^i + \epsilon_t^i, \; \epsilon_t^i \sim N(\mu_t^i, \sigma^2). \tag{2}$$

I assume that country B grows faster on average: $\mu_t^B \geq \mu_t^A$. Note that at time 1, $e^{\mu_1^B - \mu_1^A}$ is a proxy for the relative size of two economies. After year 45, growth rates are set to be equal and the two economies are assumed to be generally equal in size.

Both labor and capital are chosen before a shock at time t. This choice means that both production factors are immobile and can not be immediately reallocated to the other sector. I denote, by $k_{t-1}^{iC}, l_{t-1}^{iC}, i = \{A, B\}$, the capital and labor used in the consumption sector at time t and, by $k_{t-1}^{jI}, l_{t-1}^{jI}, i = \{A, B\}$, the capital and labor used in the investment sector at time t.

Electricity in both countries is produced using similar decreasing returns to scale technology, but with only one production factor (e.g., coal or oil):

$$A_j^i(O_t) = e^{\mu_t^{iu}} O_t^{\alpha_u}, \quad i = \{A, B\}, j = \{c, d\}, t = 1, 2, 3..., \tag{3}$$

where O_t is the amount of fuel used to produce electricity and μ_t^{iu} is the growth rate for utilities' output. Lower index $j = \{c, d\}$ denotes a cleaner or dirtier firm, respectively. The fraction of clean utilities at country i is N^i and the fraction of dirty utilities is $1 - N^i$. The growth rate is chosen in such a way that the steady state growth in the case with no caps in country $i = \{A, B\}$ is equal to μ_t^i [1].

Also, one representative fuel producer in each country provides fuel to utilities and

[1] Namely, I choose $\mu_t^{iu} = \mu_t^i(1 - \alpha_u + \alpha_R)/(1 + \alpha_R)$.

maximizes one-period profits [2]:

$$G^i_O(O_t) = p^O_t O_t - c_O O_t^{1+\alpha_O}, \quad i = \{A, B\}, j = \{c, d\}, t = 1, 2, 3..., \tag{4}$$

where p^O_t is the time-t price of fuel, O_t is the total amount of fuel extracted and c_O is the cost of extraction.

1.2.2. Consumers

In country $i = \{A, B\}$, the endowment of a representative consumer at time 0 is one share of each of the firms in country A. At each period t, the agent maximizes the expected utility from leisure and consumption, and allocates the total available time of one among leisure and working for the two firms: consumption and investment.

I borrow arguments from Kiley (2010) that the separable-habit-formation utility function best explains the U.S. consumption data. Consumer's preferences are separable in consumption and leisure and exhibit "deep habits", as in Ravn et al. (2006):

$$U(c^{iA}_t, c^{iB}_t, l^i_t) = \log(c^{iA}_t - bc^{iA}_{t-1}) + \log(c^{iB}_t - bc^{iB}_{t-1}) + \log(1 - l^i_t), i = \{A, B\},$$

where c^{ij}_t is time-t consumption of agent i of the good produced in country j, and l^i_t is the total labor supply.

I introduce labor into this model to generate business-cycle implications, including macroeconomic moments for the labor supply, consumption, investment, and output. It is well-known that the introduction of labor complicates the solution of the model and thus places a higher burden on the ability of the model to match business-cycle behavior. I assume capital market frictions to generate a not perfectly elastic capital supply and increase volatility in capital gains necessary to produce high volatility in the return on equity. Habit formation allows me to produce a high risk premium and a reasonable risk-free rate.

[2]Note that I assume no difference in the fuel that is used by electricity generators and rather model the emission as the exogenously given coefficient of a plant's technology.

1.2.3. The Cap-and-Trade System and Cleaner Technology Installment

In this section, I introduce the concept of a cap-and-trade system. Namely, it is assumed that there are caps on the emissions of an electricity-producing firm (and, in some specifications, on investment good producers). These caps are set by the government and are exogenous to the firms.

If there is a cap e_t^j on the emissions from a firm $j \in \{c, d\}$ of a country $i \in \{A, B\}$, and if the price for permits to emit more than this fixed amount is q_t, then the profit of the firm is given by

$$\pi_t^{ij} = p_t^{el,i} A_j^i(O_t) - p_t^{fuel} O_t - q_t(\theta^{ij} A_j^i(O_t) - e_t^{ij}) - \text{clean technology cost}, \qquad (5)$$

where $p_t^{el,i}$ is the price of electricity, and θ^{ij} is the coefficient capturing higher or lower initial emissions. Without loss of generality, I assume that the emissions are proportional to the total output of a firm with the constant coefficient θ^{ij} (if the firm does not install clean technology) [3]. I assume that the number of initial permits equals a fraction of the initial emissions, and I denote this fraction f. This fraction means that the firm has $f \times emission$ free permits at the initial date and the number of them available for free decreases by a factor of e^{γ_E} each period; here e^{γ_E} is an exogenously given rate determined by the government. Because there is growth in the economy, the number of permits decreases even faster in real terms absent a cleaner technology installment.

Now I introduce cleaner technology. This technology allows a firm to decrease emissions by a factor of e^{γ} forever. If the firm installs cleaner technology in some period and its previous emissions are y, then its new emissions become

$$\tilde{y} = \frac{y}{e^{\gamma}}. \qquad (6)$$

There is a cost x to install cleaner technology. In the current version I consider the

[3] The choice of the emission function only changes q_t because I may consider any increasing function of output and solve for optimal production level; this would only change price of permits and will not change qualitative results for production

cost x proportional to the total *emissions* of a firm; this cost, in general, depends on the company, country and period. Consistent with the empirical evidence, I assume that this cost (x_t^j, $j \in \{c, d\}$) decreases over time. I am ambivalent about the actual comparison of costs for dirty and clean firms and set them to be equal among electricity producers in the baseline model.[4]

I assume that firms play a Nash equilibrium for the clean technology investments. Subsection 2.6 provides a simple example of these decisions in a two-period model. I show in subsection 2.4 that the problem of the cleaner technology installment can be rewritten to become Markovian (time is an additional state variable). Moreover, the derivation shows that, in equilibrium, firms invest only if they observe a high enough price for permits (threshold strategy).

If country B joins the agreement, its endowment is the same number of emission permits as one of country A. I relax this assumption in the extension of the model to the extent that the level of permits is enough to make country B indifferent to joining or not joining the agreement. This assumption has another impact: country B becomes a net seller of permits in the early periods and a net buyer in the later because the firms in B consider it to be expensive to install in the early periods. This result is very plausible because there is greater concern and motivation in developed countries for clean technology improvement than there is in developing countries, which concentrate on economic growth.

1.2.4. The Firm's Maximization Problem

I consider competitive firms that take state prices, permit prices and wages as given.

A filtered probability space $(\Omega, F, [F_t; t = 0, 1, ...], \mathbf{P})$ represents uncertainty generated by production shocks in countries A and B. All the stochastic processes are

[4] Martin et al. (2009) find almost no difference in indirect measures for these costs in UK data; Fowlie (2010) shows that the costs in the case of NOx emissions are almost the same for the cleaner and dirtier power plants and seem to be higher for the dirtier ones. In our setting, higher costs for dirtier companies lead to even stronger results.

assumed to be adapted to $[F_t; t = 0, 1, ...]$. I assume that all processes and expectations are well-defined and satisfy necessary regularity conditions.

I assume that there are enough primitive securities to complete financial markets. The assumption on a dynamically complete market allows for easy representation of the state price density: I denote $\xi_t = \xi(\omega, t)$ the state price density at state ω and time t. Then, the value of the firm given a dividend stream is

$$V_t = \mathbf{E}\left(\sum_{s=1}^{\infty} \frac{\xi(\omega, s+t)}{\xi(\omega, t)} \pi(s+t)|F_t\right). \tag{7}$$

Firms can also trade the same securities as individuals and maximize their expected values.

Firms have to choose the timing of the clean technology investment taking into account future returns from today's change in emissions function and, thus, in the number of emission permits they are able to sell. They can decide on installment in each period and observe emission-cap constraints that are local for country A (in this case $\chi^B = 0$) or global ($\chi_B = 1$):

$$\frac{E^A + \chi^B E^B}{e^{s\gamma^E}} \geq \frac{\theta^{Ac} N^A e^{\mu_s^{Au}} (O_s^{Ac})^{\alpha_u}}{e^{\gamma S_s^{Ac}}} + \frac{\theta^{Ad} (1 - N^A) e^{\mu_s^{Au}} (O_s^{Ad})^{\alpha_u}}{e^{\gamma S_s^{Ad}}} +$$
$$+\chi^B \frac{\theta^{Bc} N^B e^{\mu_s^{Bu}} (O_s^{Bc})^{\alpha_u}}{e^{\gamma S_s^{Bc}}} + \chi^B \frac{\theta^{Bd} (1 - N^B) e^{\mu_s^{Bu}} (O_s^{Bd})^{\alpha_u}}{e^{\gamma S_s^{Bd}}}, j = \{A, B\},\ s = t, t+1, ... \tag{8}$$

.

Note that I can rewrite (8) in the following way:

$$E^A + \chi^B E^B \geq (1 + \gamma^E)^s \times RHS, \tag{9}$$

where RHS only depends on the variables at time s. Last inequality shows that the problem is Markovian in the sense that its solution at time t only depends on the variables at time $t-1$, including S_{t-1}^{ij}. However, it is not stationary because the solution depends on time as the growth rate is decreasing.

Consider first the case in which country B does not join the cap-and-trade system. In this case, electricity producers compare four possible scenarios: (1) nobody installs,

(2) only a cleaner firm installs, (3) only a dirtier firm installs, and (4) both firms install. Payoffs of the firm Ac are summarized below: if nobody installs, the firm maximizes

$$p_s^{el,1}\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u} - p_s^{fuel,1}N_AO_s - q_s^1\left(\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u}e^{-\gamma S_s^{Ac}} - e_s^{Ac}\right) + \beta V_1^c, \tag{10}$$

where V_1^c is the expected value of the cleaner firm if no installment is done in this period and $p_s^{el,1}, p_s^{fuel,1}, q_s^1$ are the prices of electricity, fuel and permits. If only firm Ac installs, then it maximizes

$$p_s^{el,2}\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u} - p_s^{fuel,2}N_AO_s - q_s^2\left(\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u}e^{-\gamma(S_s^{Ac}+1)} - e_s^{Ac}\right) - \\ -x_s\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u}e^{-\gamma S_s^{Ac}} + \beta V_2^c, \tag{11}$$

and similarly for a dirtier firm installment, and for installment by both firms installment I have

$$p_s^{el,3}\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u} - p_s^{fuel,3}N_AO_s - q_s^3\left(\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u}e^{-\gamma S_s^{Ac}} - e_s^{Ac}\right) + \beta V_3^c, \tag{12}$$

$$p_s^{el,4}\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u} - p_s^{fuel,4}N_AO_s - q_s^4\left(\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u}e^{-\gamma(S_s^{Ac}+1)} - e_s^{Ac}\right) - \\ -x_s\theta^{Ac}N^Ae^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u}e^{-\gamma S_s^{Ac}} + \beta V_4^c. \tag{13}$$

To address the optimal installment policy, I numerically optimize the choice for the firms assuming that the expected values $V_i^c, i = 1, ..., 4$ are the exponential functions of the state variables (namely, the ones I describe in the Appendices 1 and 2). I show that all four outcomes may be observed in the model.

In the case when both countries participate in the system, I have even a greater number of choices (namely, eight). Yet again, it is relatively easy to solve all eight possibilities assuming that the future values of the firms in the electricity sector are given by the exponential polynomials of the state variables.

1.2.5. Equilibrium of the Economy

In complete markets it is possible to solve for competitive equilibrium using a central planner approach. The central planner maximizes the weighted expected utility of two representative consumers:

$$\max \sum_{s=t}^{\infty} \beta^{s-t} \{ \log(c_s^{AA} - bc_{s-1}^{AA}) + \log(c_s^{AB} - bc_{s-1}^{AB}) + \log(1 - l_t^A) +$$
$$+ z \log(c_s^{BA} - bc_{s-1}^{BA}) + z \log(c_s^{BB} - bc_{s-1}^{BB}) + z \log(1 - l_t^B) \} \quad (14)$$

given $c_{t-1}^{ij}, k_{t-1}^{iC}, k_{t-1}^{iI}, l_{t-1}^{iC}, l_{t-1}^{iI}, i = \{A, B\}, j = \{A, B\}$ and budget constraints:

$$(a_s^{iC})^{\psi} (Z_s^i l_{s-1}^{iC})^{(1-\alpha)(1-\psi)} (k_{s-1}^{iC})^{\alpha(1-\psi)} \geq c_s^{Ai} + c_s^{Bi},\ i = \{A, B\},\ s = t, t+1, ..., \quad (15)$$

$$(a_s^{jI})^{\psi} (Z_s^j l_{s-1}^{jI})^{(1-\alpha)(1-\psi)} (k_{s-1}^{jI})^{\alpha(1-\psi)} + (1-\delta)(k_{s-1}^{jI} + k_{s-1}^{jC}) \geq k_s^{jI} + k_s^{jC},$$
$$j = \{A, B\},\ s = t, t+1, ..., \quad (16)$$

$$N^j e^{\mu_s^{ju}} (O_s^{jc})^{\alpha_u} + (1 - N^j) e^{\mu_s^{ju}} (O_s^{jd})^{\alpha_u} \geq a_s^{jC} + a_s^{jI}, j = \{A, B\},\ s = t, t+1, ... \quad (17)$$

Furthermore, we have an emission constraint (8). In the case of non-binding caps, Lagrange multiplier of the last constraint is zero, as well as the price of permits; all the other Lagrange multipliers should be positive.

I solve the model in quarterly periods. The solution is based on the first order conditions derived from the central planner's problem and the budget constraints. I approximate expectations of the future variables with exponential polynomials of the state variables. To check the accuracy of the approximation I implement Den Haan and Marcet's (1994) test (see Appendix 1).

1.2.6. An Example of Cleaner Technology Installment and Returns

In this section, I provide a simple two-period example to highlight the reasoning behind the results. Assume that there is one representative consumer with log preferences, no time discounting, and no habits. There is one country and there are two firms (A and B) that differ in their emissions. These firms produce one final good. There is only one state at time 1 and two equally likely states at time 2. Assumptions on the outputs are below:

Firm	Date 1	Date 2, boom	Date 2, recession
A, output	1	1	0.9
B, output	1	1	0.9

Without emission caps, the stochastic discount factor is equal to $2/1.8 = 1.111$ in a recession and 1 in a boom and the values of the firms at date 1 are equal to $1 + 0.5 \times 1 + 0.45 \times 1.111 = 2$. The expected returns are -5%.

Now assume that emissions are proportional to output and that the emissions of A are 1 and those of B are 2 at the initial date. Let the government set a cap to decrease at the 10% level so that initial emissions of 3 should be reduced to 2.7 at date 2. Assume that the clean technology allows a reduction in emissions of 20%, and that the cost of its installment is equal to x. In the recession state, there is no need to reduce emissions as they are equal to $0.9 \times 1 + 0.9 \times 2 = 2.7$.

In the boom state, with no clean technology installment both firms should reduce their production to 0.9. If only firm A installs then the profits of the firms are

Firm	Output	Cap	Emission	CT spending	Price of permits	Total revenue
A	1	0.9	0.8	x	0.5	$1 + 0.5 \times 0.1 - x$
B	0.95	1.8	1.9	0	0.5	$0.95 - 0.5 \times 0.1$

The price of permits in this case is derived from the equation $0.95 - 0.1p = 0.9$ because the firm might choose between either reducing the production to 0.9 or paying the price of the permits.

I can compare this to the case in which only firm B installs:

Firm	Output	Cap	Emission	CT spending	Price of permits	Total revenue
A	1	0.9	1	0	0	1
B	1	1.8	1.6	$2x$	0	$1-2x$

Note that the price of permits is zero even if firm A has positive net emissions. Yet the reduction in firm B's emissions is so high that the firm A can still produce the same output and pay nothing for emissions as there is a net positive supply of permits.

Now compare all three scenarios [5]. Scenario "only B installs" is beneficial for firm B if

$$1-2x>0.9,\ x<0.05.$$

In this case, firm B would install and firm A will free-ride. If $x>0.05$, firm B declines to install and reduces output. But in this case, firm A is better off installing the clean technology if

$$1.05-x>0.9,\ x<0.15.$$

This means that if the cost is in the range of $0.05<x<0.15$, the optimal strategy is "only firm A installs clean technology". If $x<0.05$ then firm B installs and firm A free-rides, and if $x>0.15$, then nobody installs.

What does this mean for the returns? Take the case in which $0.05<x<0.15$, e.g. $x=0.06$. Then the following are the total revenues and returns for the firms:

Variable	Date 1	Date 2, boom	Date 2, recession	Firm value	Returns
A	1	0.99	0.9	1.0239	-7.697%
B	1	0.9	0.9	0.9762	-7.8%
SDF	1	1.0582	1.111		

The value of firm A's future payments is $0.5\times0.99\times1.0582+0.5\times0.9\times1.111=$

[5] The fourth scenario, both firms install, is inferior from the point of view of the firm A to the scenario in which only firm B installs

1.0239, and that of the firm B is $0.5 \times 0.9 \times 1.0582 + 0.5 \times 0.9 \times 1.111 = 0.9762$. The expected return for the firm A is $(0.5 \times 0.99 + 0.5 \times 0.9)/1.0239 - 1 \approx -0.077$, and for the firm B is $0.9/0.9762 - 1 \approx -0.078$. Results are pretty much the same for any $0.05 < x < 0.15$. This simple example shows that expected returns on a cleaner firm's equity are higher than these of a dirtier firm. The main ingredient of the result is a higher correlation of the returns of firm A with the SDF, but note that here I abstract from any intertemporal decisions for the firms that were introduced in subsection 2.4.

1.3. Calibration

The main parameters of the model are in Table 1.1.

I take the technology parameters $(\alpha, \sigma, \mu_t^A, \delta)$ from Boldrin et al. (2001). The ψ is slightly higher than the electricity expenditures by manufacturers in the UK in 2006 (7.79% of total expenditures) and the energy expenditures in the U.S. in 2007 (8.8% of GDP). The growth in country B starts at 6% annually, which is equal to the projected growth in BRIC countries in 2010. The assumption on the relative size of the economies $e^{\mu_0^A - \mu_0^B}$ is three, which is two times higher than in the data but this is not important in the model as it only changes relative prices. The β is set to match the risk-free rate, while b is chosen to match the risk premium and Sharpe ratio, and ν is chosen to equate labor supply to 1/3 in the steady state.

Emission cap decrease is consistent with the goal to reduce the emissions by 20% in 15 years. The γ is the coefficient for which I do not have a plausible empirical estimate; therefore, I check for the impact of this coefficient on the results. I only assume that $\gamma > \gamma_E + \mu^A$, that is, the clean technology allows firms to reduce emissions more than the (nominal) caps demand. My baseline estimate of γ is roughly equal to that of the 2006-2007 UK emission trading system[6]: in those two years, firms were able to reduce emissions by 4.44% that equals a 0.54% real reduction in each quarter. Figure 1.2 shows current emissions and GDP of the main emitters.

The parameters θ^{ij}, $i = \{A, B\}$, $j = \{c, d\}$ can be reduced to only two as I assume that all the power producers in country B have high emissions and half of the power producers in the country A have low emissions [7], i.e., $N_A = 0.5$ and $N_B = 0$. Because of that, I only need two coefficients θ^{Ac} and $\theta^{Ad} = \theta^{Bd}$. I set $\theta^{Ac} = 1$ and $\theta^{Ad} = 1.7$ that is consistent with the EIA (2005) report. The cost of a cleaner technology installment,

[6] I estimate γ on the UK data because implementation of the market for permits is better in the UK than in EU Emission Trading System in this period.

[7] In 2009, coal burners represent 44.9% of the total electricity production in the US, while natural gas, hydroelectric and nuclear power stations contribute 50.6%. In China in 2009 total production was mostly divided between coal plants (83%) and hydropower (14%).

$x_t = 0.1$, is set to decrease with the rate of 0.001 per year and is chosen in a way to ensure that the installment of the clean technology is done first by a cleaner firm in country A [8]. Finally, the cost of extraction $c_O = 0.005$ is chosen to produce the interior solution for fuel extraction.

In what follows I will consider four 15-year phases of development; these phases differ due to country B's growth rate as this growth rate is continuously decreasing over time. In the last phase (phase 4) both countries grow at the same average rate, $\mu_t^A = 0.04\%$, which equals to approximately a 1.61% annual rate. In phase 1, country B generally grows at the rate 5.26% per year; in phase 2, the growth rate reduces to an average of 3.8%, and in phase 3, it reduces to 2.33%. I also assume that during phase 4 the emission caps stay the same for every period and clean technology costs nothing. Thus, the firms install clean technology only when they expect the price of emissions in the next period to be positive, but they do not pay for this and I can solve for a balanced growth path.

1.4. Results

I consider several scenarios in which country B joins the cap-and-trade agreement. Note that there are two dimensions of the choice: (1) does country B join, and when, (2) do manufacturing plants (which I represent by an investment good producer) participate in the system? I assume that utilities are always included in the system (if a country participates in the agreement), yet manufacturing plants might not[9]. However, this is not a plant's choice to participate - rather, I solve different scenarios in which an investment good producer either participates from the time of a country's signing the agreement, or it never participates. In the current version, I only solve

[8] This assumption is consistent with the evidence presented in Fowlie (2010). See also the example of the subsection 2.6.

[9] Electricity producers in the U.S. account for more than 40% of the man-made carbon dioxide emission and thus they represent the most important part of the regulation. For the UK, this number is more than 46%.

the scenarios without an investment good producer included in the system.

As the benchmark case (case 1), I solve the model with country A starting a cap-and-trade system within the country, country B staying outside the system, and with only utilities involved in permit trading. I call this case 1 (Developed Country Agreement Without Manufacturing Plants).

Case 2 is the case in which country B joins the system at period 1 and only utilities are involved (Worldwide Agreement Without Manufacturing Plants).

1.4.1. Case 1 - Developed Country Agreement

In this benchmark case manufacturing plants are excluded from the cap-and-trade system, yet country A introduces the caps on total emissions to the electricity firms equal to the initial emissions at period zero. Note that country B is also indirectly impacted by the introduction of an agreement, because firms in country A decrease their production in response to the positive price of permits and cleaner technology costs, and the price of country A's consumption good increases relative to the price of country B's consumption good. Thus, the welfare of the country B might decrease in this scenario.

Below, I report the main asset pricing statistics for the baseline case (see Tables $1.2 - 1.4$). We see that the cleaner technology installment decreases for a cleaner firm and increases for a dirtier firm. I also note that the number of permits traded decreases in later periods.

I can almost match the risk premium and the Sharpe ratio in the data. The main result I obtain is the consistent difference between the returns on equity for cleaner and dirtier firms. The reason for that result is the following.

First, let's start with the installment choices. The cleaner firm installs more often than the dirtier one in every phase of development, and this leads to the result that the cleaner firm is always selling permits, while the dirtier firm is always buying. Firms install simultaneously in some periods, and they each install in more than half of the periods to satisfy the caps.

Regarding the return spread, I get a similar result to the one in the simple example of subsection 2.6. The firm with lower emissions has a lower absolute cost of installing the cleaner technology and installs more often than the dirtier one. Moreover, the dirtier firm installs more when demand for electricity is high (shock to productivity in consumption and investment sectors is high). This means that the dirtier firm installs cleaner technology exactly when the stochastic discount factor is high, and it is also the buyer of permits. Thus, the cleaner firm sells permits to the dirtier one in a boom and both firms spend less on permits in a recession as the price of permits lowers, sometimes to zero (either because both firms are below their caps or because the net supply of permits is positive). All this means that the dirtier firm has cash flows and returns *less* correlated with the stochastic discount factor, and are less risky.

The model also provides business cycle implications. In Table 1.5, I report the macro variables. Note that labor supply is reasonable for both countries and decreases over time, which is consistent with the data. Moreover, consumption growth volatility decreases over time and is in line with the empirical estimate at the end.

I compare my results based on the polynomial approximations with the exact solution but I can not do so directly as there is no analytical solution. To check the accuracy of the approximation, I implement Den Haan and Marcet's (1994) test to compare residual statistics with χ^2_{13} or χ^2_4 quantiles (see Appendix 1). It appears that the approximations are quite reasonable: test statistics are within the theoretical bounds in more than 95% of the simulations for a long series of the simulated data.

The results obtained in this section are robust to the changes in parameter γ. For comparison, I report asset pricing statistics generated for country A in two other cases: when $\gamma = 0.009$ and when $\gamma = 0.0012$ (see Tables 1.6 and 1.7). Results are quantitatively similar and qualitatively the same. The only thing to mention is that the difference in returns between a cleaner firm and a dirtier firm decreases in γ. This result is intuitive: the higher the reduction in emissions is from every improvement in technology, the lower the price of permits is and the lower the correlation of clean-firm equity returns is with the SDF.

1.4.2. Case 2 - Worldwide Agreement

Consider now the case in which both countries join the emission agreement from the very beginning. I get the following results (see Tables 1.8-1.11) [10].

The trade in permits is higher in the beginning and the installment of clean technology is lower than in the benchmark case. This is the result of my assumption: I give country B a higher number of permits, and thus it immediately reduces the need to install cleaner technology and instead leaves the countries trading permits.

Why do I observe a level difference in returns between two scenarios [11]? First, adding country B to the permit traders immediately reduces the price of permits to the lowest bound for a long period of time as the initial level of permits in country B is higher than needed for its own production. This level leads to more smooth and higher profits and returns. Second, firms invest less in cleaner technology in the early periods and do not reduce their emissions until the emission market constraints become tightly binding. This trend has an interesting implication: despite the decrease over time of the relative costs of clean technology installment (x_t), absolute costs might not be so because the emissions themselves increase with the increase in production levels. Power plants find it more expensive to install cleaner technology than in the first scenario and thus returns on equity become more correlated with the stochastic discount factor.

I also observe a cross-sectional difference for returns of cleaner and dirtier electricity producers. Returns are consistently higher for the dirtier firm. Again, this is mostly the result of the need for the dirtier firm to buy permits, especially in booms.

This scenario leads to a higher expected utility for both consumers (19% higher

[10] Again, Den Haan-Marcet test shows that I have a good approximation.

[11] Note that both risk-free rate and returns on equity depend on the growth rate almost one-to-one (while not fully one-to-one because the model is not stationary). This is imminent implication of our assumptions.

than in the baseline case) as they now can consume more in the early periods. This is the implication of less strict emission caps: because of that, firms in country A produce more in the early periods. The agent discounts future consumption, and the reduction in consumption in future periods which follow from the binding caps in phases 3 and 4, does not hurt him or her.

However, this scenario with such a large number of permits given to country B (it gets 133% more permits than it should) leads to a lower reduction in emissions than the first scenario. Accumulated emissions are more than twice as high in this case, which is an undesirable result of the equilibrium.

1.4.3. Number of Permits Needed to Make Country B Indifferent

In this section, I derive the number of permits needed to make country B indifferent to joining or not joining the cap-and-trade system. To do that, I compare the overall utility derived from the consumption streams for the agent in country B in the baseline case and in the case in which country B gets a fixed amount of permits. The model is simulated 400 times in order to produce average utility for a given number of permits.

I estimate the break-even number of permits to be equal to 72% of the permits in country A. This is still two times higher than the initial emissions of country B, but even this scenario does not help to resolve the emissions issue. Namely, the accumulated emission is 46% higher than in the case with no country B in the system. This is the result of two dependent effects: First, firms face soft constraints on emissions in the early periods and do not install the clean technology. Second, this policy leads to an increase in the absolute costs of future emission reductions; because the emissions of the firms increase due to exogenous growth, it becomes more and more costly to install the clean technology. This cost increases the threshold for the installment decision and hence reduces incentives even further.

1.4.4. Auctioning some of the Permits

In this section, I consider the case in which only a fraction, $f < 1$, of the permits is given for free in each period. That is, a firm receives the number of permits, w, which cover fraction f of its initial emissions, and the government auctions the rest to the highest bidder. This auction is only needed if the number of permits the firms can trade among themselves is not enough to cover their emissions. I concentrate on the case with country A only in the cap-and-trade system to simplify our solution. The price for the permits auctioned by the government is defined by the second highest bidder. To define this price, I compare the profits of the firms in two cases: first, they can reduce their output and buy no permits; second, they can buy permits and produce as much as they need. The second highest bidder is indifferent between the two policies in this case.

I report the financial results for $f = 0.9$[12] in Table 1.12. Note that too low a value for f might lead to no solution (profits can become negative and it could be better to close the firm rather than to produce), thus I concentrate on this moderate case.

Results are as announced: returns for both cleaner and dirtier firms reduce because the firms have to pay for permits, but this impacts the dirtier firm more than the cleaner one. Because of that, returns for the cleaner firm are higher and the spread between returns is higher.

1.5. Empirical Evidence

In this section, I shortly discuss preliminary empirical results that might support my theoretical model. I start with a small sample of 10 firms working mainly in the UK and publicly traded on London Stock Exchange at least from 1997. These firms represent approximately 25% of the overall emission of firms working in the UK in 2007-2009. By now, I was only able to match this much plants with the ultimate own-

[12] Initial allocation of permits for the power producers in EU ETS was very high and usually set around 100%, but this number falls over time.

ers because of a complicated ownership structure observed in the European market.

The firms are very different in size and emissions; moreover, macroeconomic conditions impacted these firms differently. Thus, I analyze not only their returns, but also betas and realized risk-adjusted returns (alphas). I obtain monthly share prices for the firms from Datastream, and I obtain Fama-French factors (SMB and HML) from London Share Price Database. Proxy for the market portfolio is STOXX 600 index of European firms. I calculate Fama-French alphas and betas based on 24-month windows starting from 2000 to 2009, and I also calculate market-based alpha by doing the same procedure with one factor (market returns). Emissions and permit allocations are obtained from Community Independent Transaction Log (CITL).

My model predicts that we might expect three results: (1) *realized* returns and alphas are lower for the net buyers of the permits, (2) betas for these dirtier firms (net buyers) *decrease more* after the introduction of the cap-and-trade system, and (3) correlation of realized returns (and alphas) and consumption growth is *lower* for the net buyers. Result (1) is driven by the fact that risk-adjusted return might show higher exposure to the new risk factor, price of permits. Result (3) follows from the fact that in the model, inverse of consumption growth is a proxy for stochastic discount factor, if we exclude habits from preferences. Thus, this inverse of consumption growth is expected to be more negatively correlated with returns of the cleaner firms, and hence correlation of consumption growth and returns is higher for the net sellers.

I show that we observe all these patterns in the data. Note, however, that expected and realized returns are different in nature, and I do not claim that these results fully support the implications of the model, leave aside the small sample problems.

1.5.1. Relation Between Returns and the Number of Free Permits

I start with the descriptive statistics for the net buyers and the net sellers of permits. As trading of permits only started in 2007 (in 2005-2006, the number of permits available to the firms was roughly equal to the total emissions), we have 3 years/36 months of data in which a firm may be a net seller or a net buyer of the permits. Table 1.13 provides average returns and alphas for 2007-2009, as well as the difference between returns (or alphas) at the moment and five years before the date. The results are supportive for the model: returns and alphas for the net sellers of permits are higher, and they increase more (or decrease less) than these for the net buyers after 2006.

I also address the question on the relation between alphas and the number of permits available for sale for a given firm. I run the following monthly regression (errors are clustered at a stock level):

$$x_t = a_1 emission_t + a_2 PFS_t + \epsilon_t, \tag{18}$$

where x_t is either monthly market-based alpha (α_1) or Fama-French alpha (α_3), $emission_t$ is time-t emission, and PFS_t is the number of permits available for sale at time t. To smooth alpha, I take the sum of three alphas for the past 3 months. I also consider the regression (18) in which $x_t = \Delta\alpha_1 = \alpha_{1,t} - \alpha_{1,t-60}$ or $x_t = \Delta\alpha_3 = \alpha_{3,t} - \alpha_{3,t-60}$, that is, I take the difference between alpha today and its 5-year lag [13].

The results are reported in Table 1.14, and they show that both market-based alpha and Fama-French alpha increase in the number of available free permits. This is a (weak) support for the model.

1.5.2. Relation Between Betas and the Number of Free Permits

In this section, I obtain the results for the change in betas after the introduction of the cap-and-trade system for the net sellers and the net buyers. I consider the difference between beta in month t and its 2-year lag, and run (18) for market-based betas (β_1), as well as betas from Fama-French regressions ($\beta_{Mkt}, \beta_{SMB}, \beta_{HML}$). Note that my regressor is $x_t = \Delta\beta$ in each regression.

Results are reported in Table 1.15. Note that, as predicted, the difference in betas for the net sellers is higher than that for the net buyers. Similarly, for SMB and HML

[13] I do similar calculations for 2-year lag and find no difference

factors, the difference in betas for the net sellers is lower than that for the net buyers, which is again supportive of the model as the firms in the sample are large, and lower exposure to HML or SMB factors mean higher exposure to the "size" factors.

1.5.3. Relation Between Alphas and Consumption Growth

Finally, I report the relation between quarterly returns, alphas and quarterly consumption growth in the UK, distinctly for the net buyers and the net sellers. It seems that overall the correlation is higher for the net sellers, as is reported in Table 1.16. However, all the correlations are insignificantly different from zero, and this only provides a weak support for the implications of the model.

1.6. Conclusion

This chapter is devoted to the study of a two-country model of a cap-and-trade system and its influence on financial markets. I address the question of the financial impact of a cap-and-trade system on a cross-section of returns and technological improvements for individual firms in a unified framework, which is a novelty in the existing literature.

Three main results are established. First, I show that there is a consistent difference in returns on equity for "clean" and "dirty" firms: electricity producers who either have lower initial emissions have higher expected returns on equity. Second, this result is more pronounced with the decrease in the fraction of free permits given to the firms and the decrease in the improvements from the installment of cleaner technology. Third, if a developing country joins the agreement and if it receives a high enough amount of emission permits, both countries benefit in terms of economic growth and consumption. On the other hand, this policy does not lead to a decrease in overall emissions, which destroys the goal of a cap-and-trade system. I derive the level of permits needed to make country B indifferent to joining or not joining the agreement and show that even this level of permits is too high to make reductions in emissions that are at least as high as in the stand-alone case with the cap-and-trade system in country A only.

I attempt to find empirical support for the theoretical results. I show that, for a small sample of the UK firms, preliminary results might be in line with the predictions. Firms that were net buyers of permits have lower *realized* returns and risk-adjusted returns than the net sellers, and market betas for the former *decrease more* after the introduction of the cap-and-trade system. Moreover, correlation of realized returns with consumption growth is *higher* for the net sellers of permits.

1.7. Appendix: Solution of the Model

I solve the model with the central planner because markets are complete. The Lagrangian for the problem is

$$
\begin{aligned}
L = \sum_{s=t}^{\infty} \beta^{s-t} \{ & \log(c_s^{AA} - bc_{s-1}^{AA}) + \log(c_s^{AB} - bc_{s-1}^{AB}) + \log(1 - l_t^{AC} - l_t^{AI}) + \\
& + z\log(c_s^{BA} - bc_{s-1}^{BA}) + z\log(c_s^{BB} - bc_{s-1}^{BB}) + z\log(1 - l_t^{BC} - l_t^{BI}) + \\
& + \lambda_s^{AC}\left[(a_s^{AC})^{\psi}(Z_s^A l_{s-1}^{AC})^{(1-\alpha)(1-\psi)}(k_{s-1}^{AC})^{\alpha(1-\psi)} - c_s^{AA} + c_s^{BA}\right] + \\
& + \lambda_s^{BC}\left[(a_s^{BC})^{\psi}(Z_s^B l_{s-1}^{BC})^{(1-\alpha)(1-\psi)}(k_{s-1}^{BC})^{\alpha(1-\psi)} - c_s^{AB} + c_s^{BB}\right] + \\
& + \lambda_s^{AI}\left[(a_s^{AI})^{\psi}(Z_s^A l_{s-1}^{AI})^{(1-\alpha)(1-\psi)}(k_{s-1}^{AI})^{\alpha(1-\psi)} + (1-\delta)(k_{s-1}^{AI} + k_{s-1}^{AC}) - k_s^{AI} - k_s^{AC}\right] + \\
& + \lambda_s^{BI}\left[(a_s^{BI})^{\psi}(Z_s^B l_{s-1}^{BI})^{(1-\alpha)(1-\psi)}(k_{s-1}^{BI})^{\alpha(1-\psi)} + (1-\delta)(k_{s-1}^{BI} + k_{s-1}^{BC}) - k_s^{BI} - k_s^{BC}\right] + \\
& + \lambda_s^{Au}\left[N^A e^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u} + (1-N^A)e^{\mu_s^{Au}}(O_s^{Ad})^{\alpha_u} - a_s^{AC} - a_s^{AI}\right] + \\
& + \lambda_s^{Bu}\left[N^B e^{\mu_s^{Bu}}(O_s^{Bc})^{\alpha_u} + (1-N^B)e^{\mu_s^{Bu}}(O_s^{Bd})^{\alpha_u} - a_s^{BC} - a_s^{BI}\right] - \\
& - c_O(N_A O_s^{Ac} + (1-N_A)O_s^{Ad})^{1+\alpha_O} - c_O(N_B O_s^{Bc} + (1-N_B)O_s^{Bd})^{1+\alpha_O} - \\
& - x_s\left[(S_s^{Ac} - S_{s-1}^{Ac})\theta^{Ac}N^A e^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u}\frac{N^A}{e^{\gamma S_s^{Ac}}} + (S_s^{Ad} - S_{s-1}^{Ad})\theta^{Ad}(1-N^A)e^{\mu_s^{Au}}(O_s^{Ad})^{\alpha_u}\frac{1-N^A}{e^{\gamma S_s^{Ad}}}\right] - \\
& - x_s\left[(S_s^{Bc} - S_{s-1}^{Bc})\theta^{Bc}e^{\mu_s^{Bu}}(O_s^{Bc})^{\alpha_u}\frac{N^B}{e^{\gamma S_s^{Bc}}} + (S_s^{Bd} - S_{s-1}^{Bd})\theta^{Bd}e^{\mu_s^{Bu}}(O_s^{Bd})^{\alpha_u}\frac{1-N^B}{e^{\gamma S_s^{Bd}}}\right] - \\
& + \lambda_s^E[\frac{E^A + \chi^B E^B}{e^{s\gamma^E}} - \frac{\theta^{Ac}N^A e^{\mu_s^{Au}}(O_s^{Ac})^{\alpha_u}}{e^{\gamma S_s^{Ac}}} - \frac{\theta^{Ad}(1-N^A)e^{\mu_s^{Au}}(O_s^{Ad})^{\alpha_u}}{e^{\gamma S_s^{Ad}}} - \\
& - \chi^B\frac{\theta^{Bc}N^B e^{\mu_s^{Bu}}(O_s^{Bc})^{\alpha_u}}{e^{\gamma S_s^{Bc}}} - \chi^B\frac{\theta^{Bd}(1-N^B)e^{\mu_s^{Bu}}(O_s^{Bd})^{\alpha_u}}{e^{\gamma S_s^{Bd}}}].
\end{aligned}
$$

The following are the first order conditions for consumption:

$$\frac{1}{c_t^{AA} - bc_{t-1}^{AA}} - \mathbf{E}_t\left[\frac{b\beta}{c_{t+1}^{AA} - bc_t^{AA}}\right] - \lambda_t^{AC} = 0, \tag{19}$$

$$\frac{1}{c_t^{AB} - bc_{t-1}^{AB}} - \mathbf{E}_t\left[\frac{b\beta}{c_{t+1}^{AB} - bc_t^{AB}}\right] - \lambda_t^{BC} = 0, \tag{20}$$

$$\frac{z}{c_t^{BA} - bc_{t-1}^{BA}} - \mathbf{E}_t\left[\frac{zb\beta}{c_{t+1}^{BA} - bc_t^{BA}}\right] - \lambda_t^{AC} = 0, \tag{21}$$

and

$$\frac{z}{c_t^{BB} - bc_{t-1}^{BB}} - \mathbf{E}_t \left[\frac{zb\beta}{c_{t+1}^{BB} - bc_t^{BB}} \right] - \lambda_t^{AI} = 0. \tag{22}$$

The next conditions are for the labor choice:

$$-\frac{\nu}{1 - l_t^{AC} - l_t^{AI}} + (1-\alpha)(1-\psi)\mathbf{E}_t \left[\lambda_{t+1}^{AC} (K_t^{AC})^{\alpha(1-\psi)} (Z_t^A l_t^{AC})^{(1-\alpha)(1-\psi)} (a_{t+1}^{AC})^{\psi} / l_t^{AC} \right] = 0, \tag{23}$$

$$-\frac{\nu}{1 - l_t^{AC} - l_t^{AI}} + (1-\alpha)(1-\psi)\mathbf{E}_t \left[\lambda_{t+1}^{AI} (K_t^{AI})^{\alpha(1-\psi)} (Z_t^A l_t^{AI})^{(1-\alpha)(1-\psi)} (a_{t+1}^{AI})^{\psi} / l_t^{AI} \right] = 0, \tag{24}$$

$$-\frac{\nu}{1 - l_t^{BC} - l_t^{BI}} + (1-\alpha)(1-\psi)\mathbf{E}_t \left[\lambda_{t+1}^{BC} (K_t^{BC})^{\alpha(1-\psi)} (Z_t^B l_t^{BC})^{(1-\alpha)(1-\psi)} (a_{t+1}^{BC})^{\psi} / l_t^{BC} \right] = 0, \tag{25}$$

and

$$-\frac{\nu}{1 - l_t^{BC} - l_t^{BI}} + (1-\alpha)(1-\psi)\mathbf{E}_t \left[\lambda_{t+1}^{BI} (K_t^{BI})^{\alpha(1-\psi)} (Z_t^B l_t^{BI})^{(1-\alpha)(1-\psi)} (a_{t+1}^{BI})^{\psi} / l_t^{BI} \right] = 0. \tag{26}$$

The conditions for the electricity consumption are:

$$-\lambda_t^{Au} + \psi \lambda_t^{AC} (K_{t-1}^{AC})^{\alpha(1-\psi)} (Z_t^A l_{t-1}^{AC})^{(1-\alpha)(1-\psi)} (a_t^{AC})^{\psi-1} = 0, \tag{27}$$

$$-\lambda_t^{Au} + \psi \lambda_t^{AI} (K_{t-1}^{AI})^{\alpha(1-\psi)} (Z_t^A l_{t-1}^{AI})^{(1-\alpha)(1-\psi)} (a_t^{AI})^{\psi-1} = 0, \tag{28}$$

$$-\lambda_t^{Bu} + \psi \lambda_t^{BC} (K_{t-1}^{BC})^{\alpha(1-\psi)} (Z_t^B l_{t-1}^{BC})^{(1-\alpha)(1-\psi)} (a_t^{BC})^{\psi-1} = 0, \tag{29}$$

and

$$-\lambda_t^{Bu} + \psi \lambda_t^{BI} (K_{t-1}^{BI})^{\alpha(1-\psi)} (Z_t^B l_{t-1}^{BI})^{(1-\alpha)(1-\psi)} (a_t^{BI})^{\psi-1} = 0. \tag{30}$$

The conditions for the capital choice are:

$$-\lambda_t^{AI} + \beta\alpha(1-\psi)\mathbf{E}_t\lambda_{t+1}^{AC}(K_{t-1}^{AC})^{\alpha(1-\psi)-1}(Z_t^A l_{t-1}^{AC})^{(1-\alpha)(1-\psi)}(a_t^{AC})^{\psi} + \beta(1-\delta)\mathbf{E}_t\lambda_{t+1}^{AI} = 0, \tag{31}$$

$$-\lambda_t^{AI} + \beta\alpha(1-\psi)\mathbf{E}_t\lambda_{t+1}^{AI}(K_{t-1}^{AI})^{\alpha(1-\psi)-1}(Z_t^A l_{t-1}^{AI})^{(1-\alpha)(1-\psi)}(a_t^{AI})^{\psi} + \beta(1-\delta)\mathbf{E}_t\lambda_{t+1}^{AI} = 0, \tag{32}$$

$$-\lambda_t^{BI} + \beta\alpha(1-\psi)\mathbf{E}_t\lambda_{t+1}^{BC}(K_{t-1}^{BC})^{\alpha(1-\psi)-1}(Z_t^B l_{t-1}^{BC})^{(1-\alpha)(1-\psi)}(a_t^{BC})^{\psi} + \beta(1-\delta)\mathbf{E}_t\lambda_{t+1}^{BI} = 0, \tag{33}$$

and

$$-\lambda_t^{BI} + \beta\alpha(1-\psi)\mathbf{E}_t\lambda_{t+1}^{BI}(K_{t-1}^{BI})^{\alpha(1-\psi)-1}(Z_t^B l_{t-1}^{BI})^{(1-\alpha)(1-\psi)}(a_t^{BI})^{\psi} + \beta(1-\delta)\mathbf{E}_t\lambda_{t+1}^{BI} = 0. \tag{34}$$

Finally, the first order conditions for the fuel and clean technology installment are

$$\begin{aligned}\lambda_t^{Au}\alpha_u N_A e^{\mu_t^{Au}}(O_t^{Ac})^{\alpha_u-1} - x_t(S_s^{Ac} - S_{s-1}^{Ac})\theta^{Ac}\alpha_u e^{\mu_t^{Au}}(O_t^{Ac})^{\alpha_u-1}\frac{N^A}{e^{\gamma S_s^{Ac}}} - \\ -(1+\alpha_O)c_O N_A(N_A O_t^{Ac} + (1-N_A)O_t^{Ad})^{\alpha_O} - \alpha_u N_A e^{\mu_t^{Au}}(O_t^{Ac})^{\alpha_u-1}\frac{\lambda_t^E}{e^{S_s^{Ac}\gamma_E}} = 0,\end{aligned} \tag{35}$$

$$\begin{aligned}\lambda_t^{Au}\alpha_u(1-N_A)e^{\mu_t^{Au}}(O_t^{Ad})^{\alpha_u-1} - x_t(S_s^{Ad} - S_{s-1}^{Ad})\theta^{Ad}\alpha_u e^{\mu_t^{Au}}(O_t^{Ad})^{\alpha_u-1}\frac{1-N^A}{e^{\gamma S_s^{Ad}}} - \\ -(1+\alpha_O)c_O(1-N_A)(N_A O_t^{Ac} + (1-N_A)O_t^{Ad})^{\alpha_O} - \alpha_u(1-N_A)e^{\mu_t^{Au}}(O_t^{Ad})^{\alpha_u-1}\frac{\lambda_t^E}{e^{S_s^{Ad}\gamma_E}} = 0,\end{aligned} \tag{36}$$

$$\begin{aligned}\lambda_t^{Bu}\alpha_u N_B e^{\mu_t^{Bu}}(O_t^{Bc})^{\alpha_u-1} - x_t(S_s^{Bc} - S_{s-1}^{Bc})\theta^{Bc}\alpha_u e^{\mu_t^{Bu}}(O_t^{Bc})^{\alpha_u-1}\frac{N^B}{e^{\gamma S_s^{Bc}}} - \\ -(1+\alpha_O)c_O N_B(N_B O_t^{Bc} + (1-N_B)O_t^{Bd})^{\alpha_O} - \alpha_u N_B e^{\mu_t^{Bu}}(O_t^{Bc})^{\alpha_u-1}\frac{\lambda_t^E}{e^{S_s^{Bc}\gamma_E}} = 0,\end{aligned} \tag{37}$$

$$\begin{aligned}\lambda_t^{Bu}\alpha_u(1-N_B)e^{\mu_t^{Bu}}(O_t^{Bd})^{\alpha_u-1} - x_t(S_s^{Bd} - S_{s-1}^{Bd})\theta^{Bd}\alpha_u e^{\mu_t^{Bu}}(O_t^{Bd})^{\alpha_u-1}\frac{1-N^B}{e^{\gamma S_s^{Bd}}} - \\ -(1+\alpha_O)c_O(1-N_B)(N_B O_t^{Bc} + (1-N_B)O_t^{Bd})^{\alpha_O} - \alpha_u(1-N_B)e^{\mu_t^{Bu}}(O_t^{Bd})^{\alpha_u-1}\frac{\lambda_t^E}{e^{S_s^{Bd}\gamma_E}} = 0,\end{aligned} \tag{38}$$

I also have seven budget constraints $(15)-(8)$. The assumption is that firms play threshold strategies (they only install if the price is higher than some endogenous threshold).

To solve the problem, I approximate conditional expectations in respective equations with an exponential polynomial of the state variables,

$$\exp[\phi' \times F(K_{t-1}^{AC}, K_{t-1}^{AI}, K_{t-1}^{BC}, K_{t-1}^{BI}, Z_t^A, Z_t^B, S_t^{Ac}, S_t^{Ad}, S_t^{Bd}, t, Res_A, Res_B)].$$

I follow Faraglia et al. (2010) to reduce the number of state variables (which is equal to 17 because I have to include 4 consumptions, 4 capitals, 2 productivity, 4 labors and 3 states of the clean technology installment S_t^{ij}). To do that, I start solving the model from the last phase in which S_t^{ij} are all irrelevant and approximate the expectations relevant for country A with the variables of country A only $K_{t-1}^{AC}, K_{t-1}^{AI}, Z_t^A$, and similarly for country B. Then, after converging the model with linear functions of these variables, I define the optimal residual function by an OLS regression of the rest of the of state variables on the given three. These residuals approximate the expectation and I recompute the coefficients. Furthermore, I proceed to the second moments of these four state variables and generate one optimal combination of the second moments. I end up with 5 state variables in the last period (Res_A is a 2×1 vector of the best linear combination and the best second moments combination in country A, Res_B is the same for B).

Moreover, I should approximate the value function for a value of an electricity firm (see (10)-(13)). I proceed similarly as I do to the approximation of first order conditions and assume that these values only depend on the state variables and choose exponential form for it. This is a simplification of my solution and allows me to solve the model relatively easily, while preserving the tractability.

For the other periods, I do the same but add all the other state variables mentioned in F. Doing as I did for the last phase, I end up with 14 state variables that produce a very accurate approximation for the conditional expectation.

To check that the expectations for the approximation are accurate, I implement the test of Den Haan-Marcet (1994) for the series of 400 observations of the last phase and the 180 observations of the first three phases (note that I have a non-stationary model except for the last phase and thus have to be careful that the number of periods in the model coincide with the number of periods used for the simulations). In both of the models I solve in this chapter, the approximation is very accurate as the simulated chi-squared statistics fall into the rejection area in only 3.96% of the simulations in the model with only country A introducing the cap-and-trade system and 4.22% of the simulations in the cases when both countries participate in the system.

I also use the homotopy approach to solve the last period's model. Namely, I start from the case and slowly increase to the level of 0.7. This is done to simplify the initial system I solve because the approximation for the case with is hard to obtain if I start with coefficients that are far from the true ones. The result is an approximation that becomes locally unstable.

1.8. Appendix: Optimal Threshold Strategies for the Firms

In this appendix, I prove that there exists an equilibrium in which electricity producers play threshold strategies, that is, they only install clean technology if the observed price is higher than some endogenous threshold. To prove that, I start from the observation that the choice of the cost of the installment guarantee in which the clean firm installs first and for the lower permit price: this is similar to what I had in the subsection 2.4.

Now consider the first case (only country A introduces the emission cap) and turn to equations (35)-(36). Note that production increases in the level of the electricity shadow price λ_t^{Au} and decreases in the level of the shadow price of permits λ_t^E.

I compare four options mentioned in subsection 2.4 to provide the evidence that any of four cases might realize in the model. Note, however, that this result is only descriptive because the thresholds I derive are endogenous and numerically calculated in the model; I do not obtain an analytical form for them.

In each scenario, firms can recalculate the prices for the fuel and permits, and thus compute the current and future value of the firm in each policy. I show next that (1) both firms install if $q_s^1 > p_1, q_s^2 > p_2$, (2) only dirtier firm installs if $q_s^1 > p_1, q_s^2 \leq p_2$, (3) only cleaner firm installs if $q_s^1 \leq p_1, q_s^2 > p_2$, and (4) nobody installs if $q_t^1 \leq p_1, q_t^2 \leq p_2$. Both thresholds are endogenous and depend on the state variables.

Compare (10) and (12) first. After the dirtier firm installs the technology, the prices of permits and of electricity decrease (because both firms might produce more), while the price of fuel increases (as demand for it increases). Moreover, the cleaner firm expects to sell less permits in future periods. This expectation leads to a decrease in the value of the cleaner firm with respect to option (10), and it has to compare this to option (13). In this option, the price of electricity falls even more, the price of fuel increases with respect to (12), and the price of permits decrease but cleaner firm is still a seller of permits, while in option (12) it may become a buyer. Thus, clean firm may either install simultaneously with the dirty firm, or allow the latter to install.

When is installment beneficial for the dirtier firm? This happens if (1) the option (12) leads to a higher firm value than the option (10) and the cleaner firm does not install or (2) the cleaner firm installs and (13) is better than (11) for the dirty firm. But note that $V_3^d > V_1^d$ and $V_4^d > V_2^d$ because the dirtier firm has to buy less permits in the future, and thus the price for them is lower. Hence, comparison of (10) and (12), (11) and (13) and the condition that the dirtier firm might only be a buyer of permits leads me to conclude that dirtier firm installs iff

$$q_s^1 \geq p_1, q_s^2 \geq p_2. \tag{39}$$

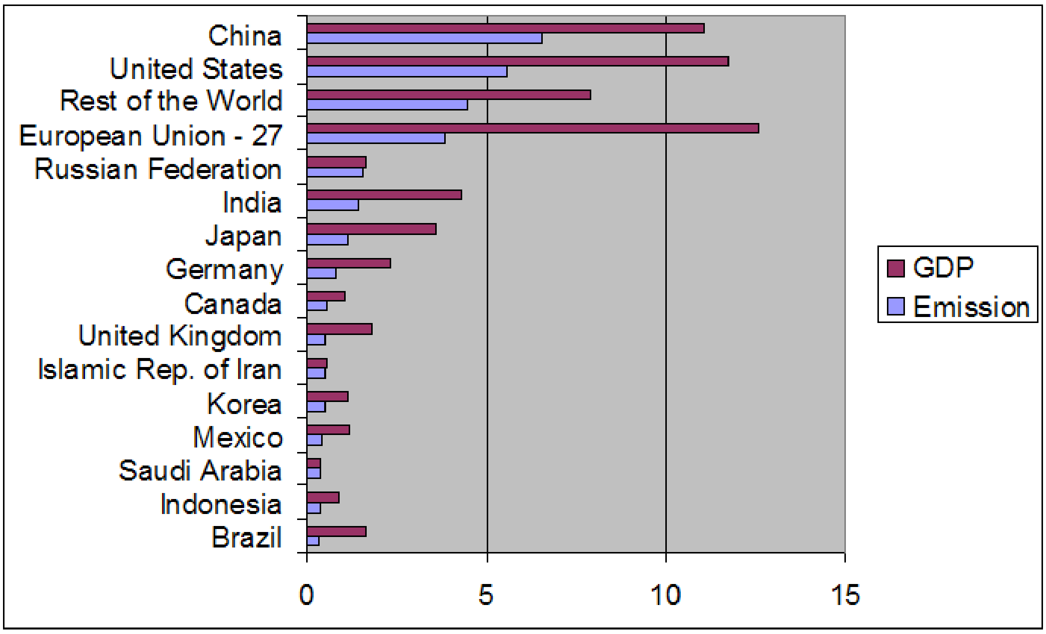
Similarly, we get other results.

1.9. Figures

Figure 1.1: EU ETS emissions allowance prices: April 2005 - December 2009

Data source: European Climate Exchange (2009). This graph shows the price changes for permits in EU Emission Trading System market.

Figure 1.2: Emission and GDP in 2008



Data source: International Energy Agency (2009). This graph shows emission (in billion tonnes of CO_2) and GDP (in trillion U.S. dollars) in the end of 2008 for the most polluting countries.

1.10. Tables

Table 1.1: Parameters of the model.

Variable	Description	Value
μ_t^A	Growth rate in A	0.004
σ	Volatility of productivity shock	0.018
α	Production coefficient	0.33
δ	Capital depreciation	0.021
ψ	Share of electricity	0.1
μ_t^B	Initial growth rate in B	0.0146
$e^{\mu_0^A - \mu_0^B}$	Time 0 relative size of A to B	3
β	Time discount	0.99
ν	Leisure utility coefficient	3.19
b	Habit formation coefficient	0.74
γ	Clean technology coefficient	0.01
γ^E	Emission cap decrease coefficient	0.0037
f	Fraction of permits given for free	1

In this table, I report the parameters for the model. For the discussion, see section 1.3.

Table 1.2: Cleaner technology installment, permits trading for case 1, country A.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**
Installment periods, Clean	53.26	50.19	47.91
Installment periods, Dirty	42.33	44.11	45.85
Number of permits traded	0.0423	0.0414	0.0277

This table reports statistics for the case with a cap-and-trade system in country A only. The first and second rows provide the average number of quarterly periods (out of 60) in which the cleaner technology is installed. The third row contains the average number of permits sold by the clean firm (initial number of permits is 2.7). For the discussion, see subsection 1.4.1.

Table 1.3: Financial statistics for case 1, country A.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**	**After year 45**	**Data**
$R_{c,i}$	6.99%	6.70%	6.41%	6.17%	6.47%
$\sigma_{c,i}$	24.38%	23.76%	22.49%	20.44%	19.64%
R_{clean}	4.88%	3.97%	3.20%	3.00%	
$\sigma(R_{clean})$	10.75%	9.13%	9.10%	9.01%	
R_{dirty}	4.71%	3.87%	3.03%	3.00%	
$\sigma(R_{dirty})$	10.30%	9.11%	9.08%	9.01%	
r_f	3.31%	2.42%	1.66%	0.96%	0.96%
σ_{r_f}	27.77%	26.92%	24.41%	23.34%	

This table reports the statistics for country A in the case with a cap-and-trade system in country A only. Here, $R_{C,I}$ is the return on consumption and investment good stocks and $\sigma_{C,I}$ is its volatility. For the discussion, see subsection 1.4.1.

Table 1.4: Financial statistics for case 1, country B.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**	**After year 45**	**Data**
$R_{c,i}$	7.83%	7.24%	6.49%	6.17 %	6.47%
$\sigma_{c,i}$	24.25%	23.35%	22.00%	20.44%	19.64%
R_{dirty}	4.75%	3.80%	2.92%	2.89%	
$\sigma(R_{dirty})$	9.13%	8.98%	8.91%	8.90 %	
r_f	3.31%	2.42%	1.66%	0.96%	0.96%
σ_{r_f}	27.77%	26.92%	24.41%	23.34%	

This table reports the statistics for the country B in the case with a cap-and-trade system in country A only. Here, $R_{C,I}$ is the return on consumption and investment good stocks and $\sigma_{C,I}$ is its volatility. For the discussion, see subsection 1.4.1.

Table 1.5: Macroeconomic statistics for case 1.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**	**After year 45**
Labor in A	0.341	0.336	0.334	0.333
Labor in B	0.428	0.395	0.348	0.333
σ_c	4.14%	3.67%	2.89%	1.89%

This table reports macroeconomic statistics in the case with a cap-and-trade system in country A only. For the discussion, see subsection 1.4.1.

Table 1.6: Financial statistics for case 1, $\gamma = 0.009$, country A.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**	**After year 45**	**Data**
$R_{c,i}$	7.13%	6.84%	6.68%	6.17%	6.47%
$\sigma_{c,i}$	25.08%	24.19%	23.50%	20.44%	19.64%
R_{clean}	5.13%	4.52%	4.01%	3.00%	
$\sigma(R_{clean})$	11.89%	10.57%	9.64%	9.01%	
R_{dirty}	5.00%	4.04%	3.72%	3.00%	
$\sigma(R_{dirty})$	11.56%	10.46%	9.47%	9.01%	
r_f	3.46%	2.51%	1.69%	0.96%	0.96%
σ_{r_f}	28.29%	26.99%	24.54%	23.34%	

This table reports the statistics for country A in the case with a cap-and-trade system in country A only and $\gamma = 0.009$. Here, $R_{C,I}$ is the return on consumption and investment good stocks and $\sigma_{C,I}$ is its volatility. For the discussion, see subsection 1.4.1.

Table 1.7: Financial statistics for case 1, $\gamma = 0.012$, country A.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**	**After year 45**	**Data**
$R_{c,i}$	6.92%	6.63%	6.38%	6.17%	6.47%
$\sigma_{c,i}$	23.98%	23.54%	22.34%	20.44%	19.64%
R_{clean}	4.76%	3.87%	3.14%	3.00%	
$\sigma(R_{clean})$	10.57%	9.10%	9.06%	9.01%	
R_{dirty}	4.69%	3.81%	3.01%	3.00%	
$\sigma(R_{dirty})$	10.25%	9.09%	9.06%	9.01%	
r_f	3.24%	2.32%	1.59%	0.96%	0.96%
σ_{r_f}	26.14%	25.74%	24.22%	23.34%	

This table reports the statistics for country A in the case with a cap-and-trade system in country A only and $\gamma = 0.012$. Here, $R_{C,I}$ is the return on consumption and investment good stocks and $\sigma_{C,I}$ is its volatility. For the discussion, see subsection 1.4.1.

Table 1.8: Cleaner technology installment and permits trading for case 2.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**
installment periods, A Clean	33.61	50.35	49.82
installment periods, A Dirty	30.24	44.92	46.38
installment periods, B	28.4	46.29	47.09
Number of permits sold by A Clean	-0.108	0.032	0.081
Number of permits sold by A Dirty	-0.257	-0.014	0.017
Number of permits sold by B	0.365	-0.018	-0.098

This table reports statistics for the case with a global cap-and-trade system. First, second, and third rows provide the average number of quarterly periods (out of 60) in which the cleaner technology is installed. Rows 4 to 6 contain the average number of permits sold by the firms (initial number is 2.7 for each country). For the discussion, see subsection 1.4.2.

Table 1.9: Financial statistics for case 2, country A.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**	**After year 45**	**Data**
$R_{c,i}$	7.48%	6.81%	6.53%	6.17%	6.47%
$\sigma_{c,i}$	24.44%	24.05%	23.35%	21.14%	19.64%
R_{clean}	5.11%	4.12%	3.24%	3.01%	
$\sigma(R_{clean})$	11.23%	10.11%	9.19%	9.02%	
R_{dirty}	5.02%	4.05%	3.07%	3.00%	
$\sigma(R_{dirty})$	11.37%	10.95%	9.36%	9.01%	
r_f	3.11%	2.14%	1.58%	0.96%	0.96%
σ_{r_f}	25.52%	24.66%	23.64%	23.34%	

This table reports the statistics for country A in the case with a global cap-and-trade system. Here, $R_{C,I}$ is the return on consumption and investment good stocks and $\sigma_{C,I}$ is its volatility. For the discussion, see subsection 1.4.2.

Table 1.10: Financial statistics for case 2, country B.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**	**After year 45**	**Data**
$R_{c,i}$	7.96%	7.02%	6.80%	6.17 %	6.47%
$\sigma_{c,i}$	24.92%	24.16%	23.31%	21.14%	19.64%
R_{dirty}	4.92%	4.06%	3.14%	2.89%	
$\sigma(R_{dirty})$	11.5%	10.13%	9.49%	8.90 %	
r_f	3.11%	2.14%	1.58%	0.96%	0.96%
σ_{r_f}	25.52%	24.66%	23.64%	23.34%	

This table reports the statistics for country B in the case with a global cap-and-trade system. Here, $R_{C,I}$ is the return on consumption and investment good stocks and $\sigma_{C,I}$ is its volatility. For the discussion, see subsection 1.4.2.

Table 1.11: Macroeconomic statistics for case 2.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**	**After year 45**
Labor, A	0.346	0.339	0.335	0.333
Labor, B	0.413	0.378	0.345	0.333
σ_c	3.51%	3.17%	2.74%	1.89%

This table reports the macroeconomic statistics in the case with a global cap-and-trade system. For the discussion, see subsection 1.4.2.

Table 1.12: Financial statistics for case 1, $f = 0.9$, country A.

Variable	**Years 1-15**	**Years 16-30**	**Years 31-45**	**After year 45**	**Data**
$R_{c,i}$	6.83%	6.42%	6.26%	6.17%	6.47%
$\sigma_{c,i}$	23.70%	23.39%	22.86%	21.10%	19.64%
R_{clean}	4.68%	3.72%	3.11%	3.00%	
$\sigma(R_{clean})$	10.54%	9.06%	9.05%	9.01%	
R_{dirty}	4.49%	3.56%	2.92%	3.00%	
$\sigma(R_{dirty})$	10.24%	9.02%	9.02%	9.01%	
r_f	3.17%	2.32%	1.58%	0.96%	0.96%
σ_{r_f}	25.93%	24.69%	23.55%	23.34%	

This table reports the statistics for country A in the case with a global cap-and-trade system, and $f = 0.9$. Here, $R_{C,I}$ is the return on consumption and investment good stocks and $\sigma_{C,I}$ is its volatility. For the discussion, see subsection 1.4.4.

Table 1.13: Descriptive statistics for returns and alphas, 2007-2009.

Variable	Net sellers	Net buyers	Difference
r_t	.0245	-.737	0.761
α_1	-2.42***	-3.56	1.14
α_3	-1.71**	-3.06	1.35
Δr_t	-1.2**	-5.66	4.46
$\Delta\alpha_1$	0.892	-0.478	1.37
$\Delta\alpha_3$	1.57*	-1.14	2.71*
N obs.	72	288	

This table reports the descriptive statistics (in percentage terms) for returns and alphas for net sellers and net buyers of permits in 2007-2009. r_t is time-t return on equity, α_1 is time-t market-based risk-adjusted return, while α_3 is time-t Fama-French risk-adjusted return. Δr_t, $\Delta\alpha_1$ and $\Delta\alpha_3$ are the differences between returns or alphas today and five years before today. Column 4 reports the difference between columns 2 and 3. *, ** and *** means significance at 10%, 5% and 1% level, respectively. For the discussion, see subsection 1.5.1.

Table 1.14: Relation between alphas, returns and the number of permits for sale, 2007-2009.

Variable	α_1	$\Delta\alpha_1$	α_3	$\Delta\alpha_3$
$emission_t$	0.115	0.139**	-0.092	0.29
PFS_t	1.5***	0.85***	1.53***	3.01***
N obs.	300	300	300	300
R-squared	3.92	2.11	4.96	8.37

This table reports the results of the regression (18) in 2007-2009 (in percentage terms). $emission_t$ is time-t emission of the firm, and PFS_t is time-t number of permits, available for sale. Column 2 reports coefficient for the market-based alpha. Column 3 reports coefficient for the difference between market-based alpha and its 5-year lag. Column 4 reports coefficient for the Fama-French alpha. Column 5 reports coefficient for the difference between Fama-French alpha and its 5-year lag. *, ** and *** means significance at 10%, 5% and 1% level, respectively. For the discussion, see subsection 1.5.1.

Table 1.15: Relation between betas, emissions and the number of permits for sale, years 2007-2009.

Variable	β_1	β_{Mkt}	β_{SMB}	β_{HML}
$emission_t$	1.31	0.656	2.47	-5.4**
PFS_t	13.3***	30.6***	-11.6***	-33.6***
N obs.	300	300	300	300
R-squared	11.64	58.75	26.78	27.94

This table reports the results of the regression (18) in 2007-2009 (in percentage terms). $emission_t$ is time-t emission of the firm, and PFS_t is time-t number of permits, available for sale. Column 2 reports coefficient for the difference between market-based beta and its 2-year lag. Column 3 reports coefficient for the difference between market beta in Fama-French regression and its 2-year lag. Column 4 reports coefficient for the difference between SMB beta in Fama-French regression and its 2-year lag. Column 3 reports coefficient for the difference between HML beta in Fama-French regression and its 2-year lag. *, ** and *** means significance at 10%, 5% and 1% level, respectively. For the discussion, see subsection 1.5.2.

Table 1.16: Correlation between consumption growth and alphas, 2007-2009.

Variable	Net sellers	Net buyers
α_1	0.102	-0.121
α_3	0.003	-0.03
r_t	0.162	0.014
N obs.	100	24

This table reports the correlations between alphas and consumption growth in 2007-2009 (in percentage terms). Row 2 reports correlation between the consumption growth and the market-based alpha. Row 3 reports correlation between the consumption growth and the Fama-French alpha. Row 4 reports correlation between the consumption growth and the raw return. *, ** and *** means significance at 10%, 5% and 1% level, respectively. For the discussion, see subsection 1.5.3.

2. Mutual Fund Performance, Fees and Flows

2.1. Introduction

Many studies, including several recent papers[14], have attempted to explore the determinants of fund flows and performance. While on average mutual funds deliver negative after-fees risk-adjusted return, there is a significant heterogeneity across their returns and fees. Moreover, the difference in performance is to a large extent related to fees. The paper by Gil-Bazo and Ruiz-Verdu (2009) addresses the question of cross-sectional difference in fees and performance and shows that there is a negative relation between fund's fees and before-fee risk-adjusted performance. I question this result and investigate the relation between fees, fund flows, risk-adjusted performance and its volatility.

In most of the empirical studies on mutual funds it is assumed that individual investors are able to fully diversify and hedge against idiosyncratic risks by investing in all the available mutual funds. This assumption means that investors only care about expected risk-adjusted returns, and these returns should be equal to zero for all the funds. Berk and Green (2004) show that if this is the case, there is a positive relation between fees, manager's ability and fund flows. Managers with higher abilities receive more money to invest (inflows) while low ability managers have less funds under management (outflows) and hence set lower fees. Moreover, there are two facts about the data. First, we observe time-series and cross-sectional difference in fees and volatility of risk-adjusted performance for the diversified mutual funds in the U.S. market. Second, average individual investor's holdings of mutual funds is not very high: in 2009, 42% of individual investors held 3 or less funds [15], including bond and money market funds which amount to 68% of such investment. While median number of the mutual funds held is 4, less than 2 of them, on average, are diversified equity

[14] See, for example, Gruber (1996); Carhart (1997); Sirri and Tufano (1998); French (2008); Fama and French (2008); Dong Lou (2010a)

[15] See "Profile of Mutual Fund Shareholders, 2009", Investment Company Institute.

funds. This raises a question of the exposure to a fund manager's ability born by an individual investor.

In this chapter I show that the volatility of before-fees risk-adjusted performance (alpha) has an impact on the mutual fund decisions regarding fees, and that this volatility (sigma) may be a risk factor for individual investors in a fund. I show that sigma contributes more than 3% to the explanation of the cross-sectional difference in fees and that the funds with higher alpha and sigma charge higher fees. Moreover, sigma is marginally successful in explaining the change in fees on a fund level. Alpha also contributes to the explanation of both fees change and fees level: funds with higher alpha have higher fees on average yet decrease fees more in response to higher past alpha. The latter result may seem counterintuitive, and to clarify it I investigate the relation between change in fees and future performance and show that an increase in fees leads to higher future before-fees performance. Although past performance leads to a counterintuitive decrease in fees for better performing funds, fees are to be considered as related to the future performance of the fund. Hence, fund managers may increase fees because they correctly predict better risk-adjusted returns for the fund in the future. Next, I discuss the implications for the fund flows and find that sigma has positive impact on the fund flows. In this part, I show that sigma explains fund flows better than second moments of past performance. Finally, I solve a simple model that provides the rationalization for the results.

Starting from Sirri and Tufano (1998), a huge literature about fund performance and flows has been developed. Many recent studies (e.g., Pollet and Wilson (2008), Huang et al. (2010), Lou (2010a), Reuter and Zitzewitz (2010)) explore this relation in different settings. Moreover, while there is a debate about mutual fund managers ability to pick stock and produce persistent positive risk-adjusted after-fee performance (e.g., Busse and Irvine (2006), Kosowski et al. (2006), Kasperczyk and Seru (2007), Fama and Frenach (2008)), there is not much of information about fund fees determinants. Two papers on the topic are especially worth mentioning. Gil-Bazo and Ruiz-Verdu (2009) study the relation between fees and before-fees risk-adjusted performance for well-diversified U.S. mutual funds. They find negative relation between performance and fees: funds that have higher past performance impose lower fees. They explain this relation by the clientele effect: investors in low-performing mutual funds may be less sensitive to performance and thus may not respond to higher fees. Second interesting study is Wahal and Wang (2011). Authors show that in late 1990s, mutual fund industry in the U.S. experienced a shift in competition: funds that started to operate after 1998 face higher competition against incumbent funds if their holdings were close enough to each other, and these competing funds started to reduce management fees, produce lower alpha and attract less inflows.

My study is also motivated by the stream of literature that investigates concentration of investments and performance of mutual funds, as well as predictability of performance based on second moments of returns. Kacperczyk et al. (2005) study industry-driven performance. In their paper they show that managers deciding to invest only in a few industries show higher performance after controlling for risk and style differences. Bali et al. (2005) reconsider the results of Goyal and Santa-Clara (2003) and show that the average stock volatility does not predict returns except for Nasdaq stocks and equally-weighted portfolio, in which case the result is driven by small stocks. They do not find significant relation between average stock volatility and future returns for NYSE/AMEX or NYSE. Cremers and Petajisto (2009) show that funds with higher deviation from appropriate benchmark produce higher alpha on average, and that funds with higher tracking error have higher fees on average. Note that they use tracking error with respect to a benchmark endogenous to a mutual fund and based on its holdings, while I concentrate on 4-factor alpha. Moreover, they do not study the relation between fees and sigma and only report descriptive statistics for their sample. My aim is to establish quantitative results and to show that there exists economically and statistically significant impact of sigma on a number of variables.

The structure of this chapter is the following. In section 2.2 I describe the data and the variables I use in the analysis. Section 2.3 is devoted to the empirical results. The results are rationalized in a simple model in the section 2.4. Section 2.5 concludes.

2.2. Data and Sample Statistics

2.2.1. Data Source

The data is obtained from the Center for Research in Security Prices (CRSP) Survivor-Bias Free U.S. Mutual Fund Database for the period from January 1980 to December 2008 (see Carhart (1997) and Carhart et al. (2002) for the discussion of this data set). I restrict the funds in our sample to be open-end diversified domestic equity mutual funds. To do this, I remove money market, bond and income, sectoral, institutional or speciality mutual funds by investigating the investment objective of a fund [16]. I remove the funds that do not provide information on expenses or have zero expenses. I also delete the funds that have less than 36 months of observations or have less than $10 million under management at any point of life of the fund. Finally, I delete observations with missing or zero total net asset. Remaining observations are cleaned to get rid of outliers in returns and expenses.

I obtain three Fama-French factors, momentum factor and risk-free rate from Kenneth French's website [17].

After this procedure, I end up with 2047 diversified domestic U.S. mutual funds. There are 266,289 fund-month observations in 348 months.

[16] Namely, I follow Huang et al. (2010) and select funds with the following Lipper objectives: CA, CG, CS, EI, FS, G, GI, H, ID, LCCE, LCGE, LCVE, MC, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, MR, NR, S, SCCE, SCGE, SCVE, SG, SP, TK, TL, UT. If there is no Lipper objective, I choose the following Strategic Insights objectives: AGG, ENV, FIN, GMC, GRI, GRO, HLT, ING, NTR, SCG, SEC, TEC, UTI, GLD, RLE. If neither of the previous two are present, I use the Wiesenberger Fund Type Code to select funds with the following objectives: G, G-I, G-S, GCI, IEQ, ENR, FIN, GRI, HLT, LTG, MCG, SCG, TCH, UTL, GPM. Finally, if a fund holds more than 80% of its value in common shares and is not mentioned in the previous categories, then it is included. Index funds are excluded based on their names.

[17] mba.tuck.dartmouth.edu/pages/faculty/ken.french/

2.2.2. Risk-Adjusted Fund Performance and Sample Statistics

I use the Carhart's (1997) four-factor model to estimate risk-adjusted performance of a fund:

$$r_{jt} = \alpha_j + \beta_{Mkt,j} Mkt_t + \beta_{SMB,j} SMB_t + \beta_{HML,j} HML_t + \beta_{Mom,j} Mom_t + \epsilon_{it}. \quad (40)$$

Here r_{jt} is fund j's before-fees return at time t in excess of the 30-day Treasury bill rate[18], Mkt_t is the market excess return, SMB_t and HML_t are two other Fama-French factors and Mom_t is the momentum factor. Term α_j is the risk-adjusted return of the fund and β's are the loadings of the fund on different risk factors.

I follow the same procedure as in Carhart (1997) and construct α in two steps. First, I estimate 40 on a rolling window of the past 36 months and obtain βs. Second, I generate $\alpha_{jt} = r_{jt} - \beta_{Mkt,j,t-1} Mkt_t + \beta_{SMB,j,t-1} SMB_t + \beta_{HML,j,t-1} HML_t + \beta_{Mom,j,t-1} Mom_t$. For the first 36 months of the fund's life, I estimate α using estimates of βs from the first 36 months. Results below do not depend on this choice, and I only report results for the whole time period. Throughout the paper, I will use the term "alpha" for this calculated 4-factor risk-adjusted before-fees excess return and term "sigma" for its standard deviation over the last 12 months.

Our sample has the following properties: mean unadjusted before-fees returns equal to 7.78% per year, risk-adjusted before-fees performance is only 2.56 bp above zero, and after-fees performance is -117 bp.

In what follows, I will always consider the regressions with time fixed effects. Monthly regressions include all time dummies for the months, and non-monthly regressions include all time dummies required for these regressions. Error terms are clustered at a fund level and this is not specially mentioned later. In what follows, regressors are always checked for pairwise correlations and only these that are not

[18] As fund returns are reported after expenses, I add total expenses to get before-fees returns. Total expense ratio is only available on annual (before 1991) or quarterly basis (after 1991). Most of the funds only change expense ratio once a year or less frequently, and I approximate monthly fees by dividing annual (or annualized quarterly) expense ratio by 12.

highly correlated are retained in the regression [19]. Where these correlations are too high, I specifically mention and discuss it.

2.3. Empirical Results

2.3.1. Alpha, Sigma and Fund Size

I start the empirical study by establishing the basic relation between alpha, sigma and the size of the fund. As is well known (see e.g. Reuter and Zitzewitz (2010)), fund size impacts the returns: the bigger the fund, the lower the returns. I aim to reestablish this result for my sample and also add the "volatility effect": when a fund increases in size, the volatility of alpha decreases. Both of these facts can be explained by the manager's decision to move her holdings closer to the benchmark because of lack of abilities, as in Berk and Green (2004). Making the portfolio more passive makes both volatility and return lower. Note also that Pollet and Wilson (2008) and Lou (2010a) establish the fact that funds indeed diversify their investments when they grow, but this may lead to a better performance, especially for small-cap funds. However, I consider well-diversified funds for which there is no positive effect from diversification.

I investigate the relation by running the following regressions:

$$\Delta\alpha_t = const + \beta_{tna,2} tna_{t-2} + \beta_{tna,diff} \Delta tna_{t-1} + age_{t-1} + \epsilon_t, \tag{41}$$

$$\alpha_{t,t-2} - \alpha_{t-3,t-5} = const + \beta_{tna,4} tna_{t-4} + \beta_{tna,diff} \Delta tna_{t-3} + age_{t-4} + \epsilon_t, \tag{42}$$

$$\alpha_{t,t-11} - \alpha_{t-12,t-23} = const + \beta_{tna,24} tna_{t-24} + \beta_{tna,diff} (tna_{t-12} - tna_{t-24}) + age_{t-12} + \epsilon_t, \tag{43}$$

$$\sigma_t - \sigma_{t-12} = const + \beta_\sigma \sigma_{t-12} + \beta_{tna,24} tna_{t-24} + \beta_{tna,diff,12} (tna_{t-12} - tna_{t-24}) + age_{t-12} + \epsilon_t, \tag{44}$$

[19] Namely, highest correlation observed is the one between age and the size of funds and equals 0.35; other pairwise correlations are lower and usually do not exceed 0.1.

where $\Delta y_t = y_t - y_{t-1}$, $\alpha_{t-s_1,t-s_2}$ is the average alpha between time $t - s_1$ and $t - s_2$, σ_t is the standard deviation of α_t during the months $t - 11$-t, tna_t is the logarithm of the total net assets of the fund (which I define as the total net asset value) and age_t is the age of the fund at time t (calculated as the number of months the fund operates [20]). First three regressions represent relation between monthly, quarterly or yearly changes in performance with respect to age, starting total net assets and incremental net assets collected before the current period. Regression (44) represents relation between change in sigma, fund size and incremental change in fund size between months $t-24$ and $t-12$. I add σ_{t-12} as a control variable because σ_t is a persistent process[21] Results of these regressions are summarized in Table 2.1.

I find strong support for negative coefficients $\beta_{tna,diff}$ in regressions $(41-43)$ and no significance for coefficient $\beta_{tna,diff,12}$ in (44). One reason I cannot find significant impact of $tna_{t-12} - tna_{t-24}$ in regression (44) is that alpha, and thus its volatility, depends on all the flows coming into the fund between periods $t - 12$ and t. This complicates a lot our goal to disentangle changes in sigma due to an increased or decreased fund size at time $t - 12$ from the changes in sigma due to an increase or a decrease in the fund size in any other period. I cannot find a better test to disentangle the influence of Δtna_{t-12}. Note, however, that both $\beta_{tna,24}$ and $\beta_{tna,diff,12}$ in (44) are negative.

The results of this section provide some support for decreasing returns to scale: higher fund size, on average, means lower alpha, and alpha decreases for a given fund when its size increases. Hence, functions I choose in the section 2.4.3 might be considered as a plausible first-order approximation for alpha and sigma as the functions of the fund size.

[20] Results do not change if we assume that age is equal to the number of years instead.

[21] I also redo the same regressions and restrict the age of a fund to be higher than 36; in this case, results are marginally different, and I do not report them.

2.3.2. Level of Fees Depends Positively on Alpha and Sigma

In this section, I show that the level of fees depends positively on both alpha and sigma. I assume that fees depend on the fund characteristics such as size, age, and turnover ratio. It is also assumed that the fees depend on the average alpha over the last year (from month $t-12$ to month $t-1$), and on sigma.

I use total expense ratio as a proxy for the fees. Later I may add front- and end-loads of a fund to the fees. However, I prefer to concentrate on the total expenses of a fund because this is a clear measure that does not demand us to guess the holding period for the investors.

To investigate the relation, I consider the following regression:

$$fees_t = const + \beta_\alpha \alpha_{t-1,t-12} + \beta_\sigma \sigma_{t-1} + \beta_{tna} tna_{t-1} + \beta_{age} age_{t-1} + \beta_{turn} turn_t + \epsilon_t. \tag{45}$$

Here $turn_t$ is the turnover ratio of the fund over the year from month $t-11$ to t. This regression combines annual and quarterly data: we can only access turnover ratio at annual or quarterly frequency, yet I keep monthly notation to simplify the comparison with other regressions.

The results for this regression are reported in Table 2.2, columns 1-2. I obtain positive coefficients of both past realized alpha and sigma. While alpha has only a small impact on explaining the dependent variable (about 0.5% of its variability), sigma has much higher impact (more than 3%) [22].

It is possible that sigma is not the best predictor among all the variables related to the second moment of alpha or returns. For example, volatility of monthly after-fee or before-fee returns may impact fees more. I have tried to replace sigma with these two measures of volatility [23] and find that they have smaller impact: they explain less

[22] Adding restriction on age does not change the results qualitatively, and I do not report them.

[23] It might be better not to use both sigma and volatility of the monthly returns in the same regression because their correlation is approximately 55%. However, using both in the same regression keeps them significant, and the contribution of sigma is higher.

than 1% of the variability in fees. Moreover, in order to test the short-term effects of second moments, I included square of past month's alpha in the regression, with and without sigma, and in both cases this regressor was insignificant at 10% level.

There is another possibility: agents react more to upside in alpha than to downside, and thus the second moment of alpha over the past year may have to be divided into positive and negative parts. I consider the following regression:

$$fees_t = const + \beta_\alpha \alpha_{t-1,t-12} + \beta_+ \alpha^2_{+,t-1,t-12} + \beta_- \alpha^2_{-,t-1,t-12} + \\ + \beta_{tna} tna_{t-1} + \beta_{age} age_{t-1} + \beta_{turn} turn_t + \epsilon_t. \tag{46}$$

Here $\alpha^2_{+,t-1,t-12}$ and $\alpha^2_{-,t-1,t-12}$ are two measures of square terms: first is the average of squares of alpha over the last year when alpha was positive, and second is the average of squares of alpha over the last year when alpha was negative, with a minus sign (so that this latter average is negative or zero).

I report the results in Table 2.2, columns 3-4. Both square terms have positive coefficients as expected. However, these two measures may not be the best explanatory variables for the level of fees. Indeed, if we consider the explanatory power of sigma and these two square terms, we observe that sigma contributes more than 3% in R-squared, and square terms explain less than 1%. This may mean that sigma is a more appropriate measure in this case.

The result I obtain is reasonable. Assume that volatility is a characteristic of the manager. If we think that investors are ready to bear risk[24], and they like upside more than downside[25], then this result is in line with positive relation of fees to volatility

[24] 80% of the respondents to Investment Company Institute (2009) indicated that they are ready to bear average or above risks for average or above returns, and 50% said they prefer average risks and average returns.

[25] Starting from Sirri and Tufano (1998), we observe that fund flows are more responsive to good performance and weakly responsive to bad performance of a mutual fund. This is a partial evidence that investors value upside more than downside.

for a fixed level of alpha[26]. I show in section 2.4 that a risk-neutral agent ($\gamma = 0$) may like volatility, that is, his expected utility may increase in the volatility of log-return.

I cannot reconcile the results regarding positive sign of past performance with those shown in Gil-Bazo and Ruiz-Verdu (2009). Namely, they find that fees are negatively related to past performance. While their sample is slightly bigger in size, based on marginally different criteria, has higher average returns and alphas, and they estimate alphas and betas on a 5-year window, this cannot explain huge difference in results.

2.3.3. Fund Flows Depend Positively on both Alpha and Sigma

In this section I consider relation between fund flows, alpha and sigma. The results I show differ from existing literature (e.g., Sirri and Tufano (1998)) as I get positive relation between fund flows and sigma; note, however, that I use volatility of 4-factor alpha rather than volatility of returns. To obtain evidence on the relation, I construct flow variable as:

$$flow_t = \frac{tna_t - (1 + ret_t)tna_{t-1}}{tna_{t-1}}. \tag{47}$$

Here ret_t is the before-fee return of a fund and tha_t is its total net assets at time t. Flow is the percentage change in the fund's total net assets from time $t-1$ to time t that does not come from a mechanical increase in the fund's assets because of the return on assets.

I run the following monthly regressions for monthly flows:

$$\begin{aligned} flow_t = const + \beta_{flow,-1} flow_{t-1} + \beta_{flow,-2} flow_{t-2} + \beta_{fees} fees_{t-1} + \beta_{\alpha,1}\alpha_{t-1,t-12} + \\ + \beta_{\alpha,2}\alpha_{t-1} + \beta_{\sigma}\sigma_{t-1} + \beta_{tna} tna_{t-1} + \beta_{age} age_{t-1} + \epsilon_t, \end{aligned} \tag{48}$$

[26] Gil-Bazo and Ruiz-Verdu (2009) and Cramers and Petajisto (2009) find similar positive relation in their data: the former paper shows positive relation of fees to volatility of before-fees returns, while the latter shows positive relation of fees to tracking error in their descriptive statistics.

$$flow_t = const + \beta_{flow,-1} flow_{t-1} + \beta_{flow,-2} flow_{t-2} + \beta_{fees} fees_{t-1} + \beta_{\alpha,1}\alpha_{t-1,t-12}+$$
$$+\beta_{\alpha,2}\alpha_{t-1} + \beta_{+}\alpha^2_{+,t-1,t-12} + \beta_{-}\alpha^2_{-,t-1,t-12} + \beta_{tna} tna_{t-1} + \beta_{age} age_{t-1} + \epsilon_t, \tag{49}$$

and the following monthly regressions for average quarterly change in flows:

$$flow_{t,t-2} = const + \beta_{flow} flow_{t-3,t-5} + \beta_{fees} fees_{t-3} + \beta_{\alpha,1}\alpha_{t-3,t-14}+$$
$$+\beta_{\alpha,2}\alpha_{t-3} + \beta_{\sigma}\sigma_{t-1} + \beta_{tna} tna_{t-3} + \beta_{age} age_{t-3} + \epsilon_t. \tag{50}$$

$$flow_{t,t-2} = const + \beta_{flow} flow_{t-3,t-5} + \beta_{fees} fees_{t-3} + \beta_{\alpha,1}\alpha_{t-3,t-14}+$$
$$+\beta_{\alpha,2}\alpha_{t-3} + \beta_{+}\alpha^2_{+,t-3,t-14} + \beta_{-}\alpha^2_{-,t-3,t-14} + \beta_{tna} tna_{t-3} + \beta_{age} age_{t-3} + \epsilon_t. \tag{51}$$

I control for several variables which were found to be important in explaining fund flows[27], including past fund flows ($flow_{t-1}$, $flow_{t-2}$), level of fees ($fees_{t-1}$), total net assets and age. Variable $flow_{t-s_1,t-s_2}$ represents average flow in the fund over months $t-s_1$ to $t-s_2$.

Results of the regressions (48)-(51) are summarized in Table 2.3, columns 1-4. I find that both coefficients β_α and β_σ are positive and that both alpha and sigma explain about 0.3% of the variability in fund flows. When regressions (49), (51) are considered, we see that square terms are as good in terms of explaining variability of fund flows as sigma, and that downside effect is not significant for both regressions, showing convex relation of flows to performance as in Sirri and Tufano (1998). Note, however, that in their paper they also consider the impact of volatility of after-fees returns on fund flows and find the opposite (negative) effect.

While this result is expected for alpha, it is not so for sigma. One may think that sigma is a proxy for riskiness of the fund and the agents should prevent themselves from holding the fund with high sigma. However, this seems not to be true for our dataset: agents invest more in the fund that has shown higher volatility in the past. There are two explanations for this result.

[27]However, I do not include total complex size and flows into competitors; I will include them later.

First, it may be a mechanical relation. We have some evidence that sigma decreases with the size of the fund. If flows increase more in smaller funds, higher sigma leads to higher fund flows. However, our regression shows that fund flows depend positively on the size of the fund which contradicts this explanation.

Second, it may be that the agents correctly predict what would happen in the future. If they expect that the manager i invests more of the fund into a passive portfolio when the size of a fund increases, and if they observe higher volatility of her today's portfolio than that for the manager j, they may believe that in the future manager i will still invest more in an active portfolio than the manager j and thus may have higher expected returns. This explanation is closer to the evidence provided in Cremers and Petajisto (2009): more active funds, that is, funds with higher deviation from an appropriate benchmark, produce higher alpha. I discuss this result more in theoretical framework in section 2.4.

2.3.4. Difference in Fees Depends Positively or Negatively on Sigma and Positively on Alpha

This section is devoted to time-series changes in fees and their relations to alpha, sigma and fund flows. Dataset allows me to check for such relations since there are more than 10,000 points at which a fund changed its fees.

I start with the following logit regression:

$$Dummy(\Delta fees_t) =$$
$$= \Lambda(const + \beta_{fees} fees_{t-1} + \beta_\alpha \alpha_{t-1,t-12} + \beta_\sigma \sigma_{t-1} + \beta_{tna} tna_{t-1} + \beta_{age} age_{t-1} + \epsilon_t). \tag{52}$$

Here $\Delta fees_t = fees_t - fees_{t-1}$, and $Dummy(\Delta fees_t) = 1$ if fees increase and $Dummy(\Delta fees_t) = 0$ if fees are reduced. Λ is the distribution function of the logistic distribution.

I include past level of fees to control for different attitude towards reducing fees: if fees are already high, it is reasonable to expect them to decrease and this regressor may increase the predictability of the changes in fees.

I divide the sample into two time periods: 1980-1999 and 2000-2008. While most of the changes happened in the second part of the sample, the first part consists of less funds, lower competition and, similar to Gil-Bazo and Ruiz-Verdu (2009) and Wahal and Wang (2010), may have different coefficients or signs of coefficients.

I report results in Table 2.4, columns 1-2. Note that in the first subsample, sigma has positive impact, while in the second it has negative impact. One possible explanation is that risk aversion of a fund manager changes over time, and second time period might consist of *more* risk-averse managers. I discuss it in more details in the section 2.4.

What is more important, however, is that we always observe strong and negative dependence on past realized alpha. This means that among all the funds changing fees, the funds that show lower performance in the recent past tend to increase fees, not to reduce them. This result seems to be counterintuitive: funds with lower performance shall decrease fees because their expected abilities are lower than in comparable funds.

To check the magnitude of this relation, I consider regression of the change in fees on the same set of regressors:

$$\Delta fees_t = const + \beta_{fees} fees_{t-1} + \beta_\alpha \alpha_{t-1,t-12} + \beta_\sigma \sigma_{t-1} + \beta_{tna} tna_{t-1} + \beta_{age} age_{t-1} + \epsilon_t. \tag{53}$$

Results of the regression are summarized in Table 2.4, columns 3-4. I find that β_α is negative and that β_σ is positive in the first period and negative in the second. Their impact, however, is small: sigma explains around 0.5% of variability in both subperiods, while alpha contributes around 1% in the first subperiod and almost 5% in the second subperiod.

At the same time, the level of fees for the funds changing fees have different behavior in two subsamples. We see that the level of past fees is (correctly) negatively related to the change in fees. What we may expect is that the funds increase fees if they had lower fees than comparable funds before the change, and they reduce fees if the level of fees was higher. To check for this possibility, I compare fees for the funds that change and do not change fees:

$$fees_{t-1} = const + \beta_+ Dummy_t^+ + \beta_- Dummy_t^- + \beta_\alpha \alpha_{t-1,t-12} + \beta_\sigma \sigma_{t-1} + \\ + \beta_{tna} tna_{t-1} + \beta_{age} age_{t-1} + \epsilon_t, \qquad (54)$$

$$fees_{t-1} = const + \beta_+ Dummy_t^+ \Delta fees_t + \beta_- Dummy_t^- \Delta fees_t + \beta_\alpha \alpha_{t-1,t-12} + \beta_\sigma \sigma_{t-1} + \\ + \beta_{tna} tna_{t-1} + \beta_{age} age_{t-1} + \epsilon_t, \qquad (55)$$

Here $Dummy_t^+ = 1$ if $\Delta fees_t > 0$, and 0 otherwise; $Dummy_t^- = 1$ if $\Delta fees_t < 0$, and 0 otherwise. Regressions (54)-(55) represent a test for the level of fees at time $t-1$ for the funds that change fees at time t. The first regression checks for the presence of the effect, while the second one also checks for magnitudes depending on magnitudes of the change in fees.

Results of these regressions are summarized in Table 2.5. Before 2000, funds that shrink fees had higher fees than comparable funds, and funds that increase fees had lower fees than comparable funds. The former is still true for the subsample after 1999, but the latter does not hold: funds that increase fees tend to have higher fees than comparable funds. This result is interesting because it also shows possible difference between two subperiods: in the recent years, funds tend to increase fees despite they already had higher fees than their competitors.

Negative relation between fees and past performance seems to be counterintuitive. We may consider another explanation: fees respond to the past performance, but at the same time, they are set regarding the future expected performance. Past performance may not be a good estimator for future performance, and thus we may end up with a spurious relation.

To check for this possibility, I implement the following strategy. I consider next year's performance of a fund that changes fees and compare it to this fund's performance over the current year. I expect that funds that increase fees show higher difference between alphas, and funds that reduce fees show lower difference between alphas than comparable funds. To check this, I run regression:

$$\alpha_{t,t-11} - \alpha_{t-12,t-23} = const + \beta_{fees} fees_{t-12} + \beta_{tna} tna_{t-12} + \beta_{age} age_{t-12} + \\ + \beta_{+} Dummy^{+}_{t-12} \Delta fees_{t-12} + \beta_{-} Dummy^{-}_{t-12} \Delta fees_{t-12} + \epsilon_t. \qquad (56)$$

I run this regression in three different forms. The first one includes only the funds that changed fees at time $t-12$. The second one includes all the funds; in this case, change in fees equals 0 if there was no change at time $t-12$. Finally, the third one includes only the fund-month observations of the funds at the same month when at least one change has been done. For each of these three regressions, I report results for two subsamples (before 2000 and after 1999). The results of the regressions are summarized in Table 2.6[28]. Note that fees change is *negative* when funds decrease fees, and thus positive coefficient β_{-} means that the funds that reduce fees worsen their performance in the new year with respect to the last year.

Note two important things. First, performance improves for the funds that increase fees, and it worsens for the funds that decrease fees. Second, these changes are economically significant: funds that increase fees by 1 basis point per month tend to increase alpha by approximately 1.7 basis points in 2000-2008. These results might mean that fund managers correctly predict improvement in future performance and adjust fees correspondingly. Moreover, for existing investors this result is positive: the fund that increases fees tend to make net profits for investors.

However, the results should be treated with caution: if we consider regression

$$\alpha_{t,t-11} = const + \beta_{\alpha} \alpha_{t-12,t-23} + \beta_{fees} fees_{t-12} + \beta_{tna} tna_{t-12} + \beta_{age} age_{t-12} + \\ + \beta_{+} Dummy^{+}_{t-12} \Delta fees_{t-12} + \beta_{-} Dummy^{-}_{t-12} \Delta fees_{t-12} + \epsilon_t, \qquad (57)$$

coefficients β_{+}, β_{-} are still positive for the whole sample and after 1999, yet they are negative before 2000, and none of these coefficients is significant. Thus, the result

[28] I get very similar results if I use dummy for fees change instead, but I prefer to concentrate on the dependence on magnitudes of fees change

of increasing performance seem to be driven mostly by time-series improvement of a fund's alpha rather than by a cross-sectional advantage of the funds that increase fees. Yet it is still some evidence which supports the idea that managers and investors are rational, and future performance of a fund is reflected in the manager's decision to change fees.

Finally, I consider dependence of fund flows on change in fees. Similar to Sirri and Tufano (1997), we may expect that reduction in fees leads to increasing fund flows, and increase in fees have the opposite impact. I check for this effect including regressors $Dummy_{t-1}^{+}\Delta fees_{t-1}, Dummy_{t-1}^{+}\Delta fees_{t-1}$ in the regression (48), and regressors $Dummy_{t-1}^{+}\Delta fees_{t-1}, Dummy_{t-1}^{+}\Delta fees_{t-1}$ in the regression (50). Results are reported in the Table 2.3, columns 5-6. As one can see, fund flows indeed decrease when fees increase, and the effect is economically significant: for each 1 basis point increase in fees, fund flows decrease by 2.3 basis points over the next month and by 9.6 basis points during the next quarter. Similarly, for 1 basis point fall in fees, there is 2.6 basis points increase in fund flows over the next month and 10.8 basis points increase during the next quarter. This result differs from the one in Sirri and Tufano (1998): they show that only negative change in fees has an impact on fund flows. Note that my sample is bigger and includes more fees changes.

2.4. The Model

In this section I describe the model and the implications for the optimal choice of the fees for a fund given that investors care both about alpha and its volatility. This simple model serves to show how empirical results of the previous section may be rationalized. However, I do not claim to test the model: it is provided only for the sake of completeness as an example of possible explanation for the empirical findings.

There is a continuum of agents indexed by $i \in [0, 1]$ and a continuum of active funds indexed by $j \in [0, 1]$. Agents may invest in any active fund or an index fund. Return function for the funds is specified in section 2.4.3. I assume that investors do not impact prices, and I solve partial equilibrium model.

Each agent (investor) is endowed with 1 unit of money and invests all this money in one active fund or in the index fund. I assume that the agent may only invest in one fund because of (not modeled) costs of account creation. There is a trade-off between active funds and index fund: the former delivers higher returns but imposes fees, while the latter is zero-fee but delivers lower return on average.

The model is essentially two-period. I assume that there may be any number of periods, but both agents and managers maximize one-period utilities, and I only derive implications for one-period change in fees and fund flows.

2.4.1. Funds' Returns

I assume that each fund possesses the technology of return generation. Namely, each fund j produces log-return of the form $r_j(\mu) \sim N(\alpha_j(\mu), \sigma_j^2(\mu))$ where μ is the total measure of agents who invest in this fund and α_j, σ_j^2 are two continuous and differentiable functions of the form:

$$\alpha_j(\mu) = \theta_j \alpha(\mu), \quad \sigma_j^2(\mu) = \nu_j \sigma^2(\mu). \tag{58}$$

I am agnostic about these functions for a moment and will further specify them in the section 2.4.3 but I demand that they are decreasing in μ [29].

Important assumption I make is that the expected return on any active fund is higher than the expected return on the index fund. While in the cross-section of the funds it appears that most of the funds do not have an ability to generate alpha [30], I assume that ex ante funds with negative expected excess return could not raise any money and thus will not invest. I also assume (without loss of generality) that the highest measure of agents who could invest in one given fund is 1 and that $\alpha_j(1) = \alpha_I$ where α_I is the expected log-return of the index fund. It means that the net (after fees) excess return on the given fund with respect to the index fund is negative if

[29] Some empirical evidence for this assumption is provided in the section 2.3.1.

[30] See, for example, Fama and French (2008).

$\mu = 1$, which gives interior solution for the measure of agents investing in a given fund.

Total return on a given fund can be written as:

$$R_j(\mu) = \exp(r_j(\mu)).$$

Each active fund imposes a fee c_j on investors and this fee is proportional to the final wealth of the fund [31]. Thus, the value an investor gets from the fund j is equal to:

$$\widehat{R}_j(\mu) = (1 - c_j) \exp(r_j(\mu)). \tag{59}$$

2.4.2. Investors' and Funds' Decisions

Investor i has constant relative risk aversion (CRRA) utility function from the return of fund j of the form:

$$U_i(\widehat{R}_j) = \frac{\widehat{R}_j^{1-\gamma}}{1-\gamma}. \tag{60}$$

Investor maximizes expected utility at time t based on the belief that log-return has normal distribution specified in 2.4.1[32]. Expected utility from the fund j if the measure of investors of this fund is μ_j is given by:

$$\mathbf{E}U_i(\widehat{R}_j) = \frac{(1-c_j)^{1-\gamma}}{1-\gamma} \exp((1-\gamma)\alpha_j(\mu_j) + (1-\gamma)^2\sigma_j^2(\mu_j)/2). \tag{61}$$

Investors require the same utility from all the funds and hence:

$$(1 - c_j) \exp(\alpha_j(\mu_j) - (\gamma - 1)\sigma_j^2(\mu_j)/2) = \exp(\alpha_I - (\gamma - 1)\sigma^2/2), \ \forall j \in [0, 1]. \tag{62}$$

Fund manager j has constant relative risk aversion (CRRA) utility with risk aversion coefficient γ_M and maximizes expected wealth given that the measure of investors

[31] We exclude performance-based fees because they are rare in the mutual fund industry.

[32] Note that given CRRA utility, the wealth invested is irrelevant for the pricing, that is, investing 1 unit is the same as investing any given amount W.

is μ_j:

$$\mathbf{E}U_j(\widehat{R}_j) = \frac{(c_j\mu_j)^{1-\gamma_M}}{1-\gamma_M}\mathbf{E}\exp((1-\gamma_M)r_j) =$$
$$= \frac{(c_j\mu_j)^{1-\gamma_M}}{1-\gamma_M}\exp((1-\gamma_M)\alpha_j(\mu_j) + (1-\gamma_M)^2\sigma_j^2(\mu_j)/2). \tag{63}$$

Manager takes (62) into account and maximizes his utility with respect to fees. In the next section I derive her optimal fees for some simple functions α_j, σ_j^2.

2.4.3. Return and Volatility Functions

I consider an example of the functions $\alpha_j(\mu), \sigma_j(\mu)$ which allows us to obtain a simple solution for the relation between fees and fund size. Assume that

$$\alpha_j(\mu) = \alpha_I - \nu_j \ln(\mu),\ \sigma_j^2(\mu) = \sigma_I^2 - \theta_j \ln(\mu). \tag{64}$$

These functions satisfy conditions $\alpha_j(1) = \alpha_I, \sigma_j^2 1 = \sigma_I^2$ and $\alpha_j(\mu) > \alpha_I, \sigma_j^2(\mu) > \sigma_I^2, 0 < \mu < 1$. Note that lower ν_j means lower abilities of the fund manager because in this case return function is lower. However, impact of θ_j is more complicated. If all the agents are risk-averse with coefficient $\gamma > 1$ then volatility impacts the fund's fees and flows in a negative way, yet if some agents have risk aversion lower than 1, they prefer higher volatility.

I assume that there is a learning process on ν_j, θ_j and I do not specify it here. I assume that higher ν_j represents the belief that the manager j has higher abilities, and realizations of α_j, σ_j^2 such that

$$\alpha_j(\mu_t) > \alpha_I - \nu_j(t)\log(\mu_j(t)), \sigma_j^2(\mu_t) > \sigma_I^2 - \theta_j(t)\log(\mu_j(t))$$

lead to higher next period estimates $\nu_j(t+1) > \nu_j(t), \theta_j(t+1) > \theta_j(t)$.

In this case from (62) and (64) we get:

$$\mu_j = (1-c_j)^{1/(\nu_j + \theta_j(1-\gamma)/2)}, \tag{65}$$

and from (63) the manager maximizes:

$$c_j(1-c_j)^{1/(\nu_j+\theta_j(1-\gamma)/2)} \times \mu_j^{-\nu_j+\theta_j(\gamma_M-1)/2} =$$
$$= c_j(1-c_j)^x,\ x = \frac{1+\theta_j(\gamma_M-1)/2-\nu_j}{\nu_j+\theta_j(1-\gamma)/2}. \tag{66}$$

This equation allows us to get the optimal fee for the manager:

$$\frac{1-c_j}{c_j} = \frac{1+\theta_j(\gamma_M-1)/2-\nu_j}{\nu_j+\theta_j(1-\gamma)}. \tag{67}$$

I derive the following implications:

Proposition 2.1. *Assume that $\gamma < 1$. Then, other things being equal:*

1) For a given level of σ_j^2, funds with higher α_j charge higher fees. Similarly, the fund j charges higher fees than the fund k if $\alpha_j = \alpha_k$ and $\sigma_j^2 > \sigma_k^2$;

2) Fund flows (defined as the difference between μ_j in two successive moments) increase both with ν_j and θ_j;

3) Difference between fees in two successive moments depends positively on ν_j;

4) Difference between fees in two successive moments depends negatively on θ_j iff:

$$1-\gamma \leq \nu_j(\gamma_M+\gamma-1). \tag{68}$$

Proof of proposition 2.1:

1) This result follows from the equation (62).

2) This result follows from the equation (65).

3) This result follows from the equation (67).

4) To prove this result, I start from equation (67) and rewrite:

$$x = \frac{1+\theta_j(\gamma_M-1)/2-\nu_j}{\nu_j+\theta_j(1-\gamma)/2} = \frac{1+\theta_j(\gamma_M+\gamma-2)/2}{\nu_j+\theta_j(1-\gamma)/2} - 1. \tag{69}$$

Note that the last fraction is monotone in θ_j and increases in it iff:

$$\frac{1}{\nu_j} \leq \frac{\gamma_M+\gamma-2}{1-\gamma},\ 1-\gamma \leq \nu_j(\gamma_M+\gamma-2). \tag{70}$$

If x increases in θ_j, fees decrease in it. This proves proposition 1. $\square$

Proposition 2.1 provides a rationalization for the results of the empirical section. Funds with higher alpha and higher sigma charge higher fees if $\gamma < 1$, and fund flows depend positively on past performance and sigma. Moreover, change in fees depends positively on expected performance.

Relation between change in fees and sigma is more tricky: until 2000 we observe negative relation, and after 1999 the relation is positive. This may mean that, in terms of the model, risk aversion of fund managers increases after 1999, possibly because of increased competition. Thus, it is still possible to rationalize this result in the model by increasing γ_M after 1999.

Note that, for $\gamma > 1$, predictions 1) and 2) regarding volatility impact are the opposite, and statement of part 4) holds if the sign of inequality 68 is the opposite. Yet to support the empirical results, I should assume that $\gamma < 1$.

Compare the results of this section with the results of two theoretical papers. As Berk and Green (2004) show, higher volatility of expected returns leads to lower fund flows because precision of the signal derived from realized return is lower. In their framework, returns have normal distribution, and volatility only impacts updating rule for the ability of the manager. Hugonnier and Kaniel (2008) show that for log-utility investors and log-normal returns in a dynamic setting, fund fees depend positively on the volatility of past performance. This is the result of more extreme (higher equity share) positions held by funds with higher fees. However, in their setting results are driven by the ability of a fund to invest in the riskless asset, an assumption that seems implausible for mutual funds. Moreover, there is no competition in their framework.

2.5. Conclusion

This paper is devoted to the study of the relation between fees and performance in the U.S. mutual fund industry. The paper of Gil-Bazo and Ruiz-Verdu (2009) shows that there is a negative relation between the two, and they provide a behavioral explanation of this relation. I take one step forward and investigate the relation between fees, fund flows, risk-adjusted performance ("alpha") and its volatility ("sigma"). I find several novel empirical facts.

First, the level of fees is positively related both to alpha and sigma, opposite to what is shown in Gil-Bazo and Ruiz-Verdu (2009).

Second, fund flows increase both in sigma and alpha. This result may mean that investors have low enough risk aversion and consider volatility of log-returns as a positive sign of fund manager's abilities. I also show that the funds that increase fees tend to obtain lower fund flows for the next one to three months, and the funds that decrease fees experience higher fund flows.

Third, I show that when a fund alters fees, this change depends positively on past sigma before 2000 but depends negatively starting from 2000. Intuitively, this might represent increase in risk aversion of the fund managers over time. Moreover, the change depends negatively on past alpha. This last result seems to be counterintuitive as funds with lower past performance increase fees, yet this might be related to the fund manager's ability to predict higher performance in the future. I show that increase in fees is followed by improved performance: for every 1 basis point change in fees, funds tend to increase alpha by approximately 1.7 basis points in 2000-2008. However, this only means that these funds revert to average performance after they change fees: in the cross-section, performance of the funds that change fees does not differ much from the other funds.

Finally, I provide a simple model that allows to rationalize these findings.

My future research in this setting is to consider trading strategies based on two variables: change in fees and sigma. As we see, both variables have an impact on future performance of the fund, and thus they may allow to distinguish between better and worse funds.

2.6. Tables

Table 2.1: Alpha and sigma with respect to the fund size.

Variable	Coefficient in (41)	Coefficient in (42)	Coefficient in (43)	Coefficient in (44)
$\Delta \log(tna)$	-0.932***	-2.57***	-0.56***	-4.75e-04
	(0.096)	(0.117)	(0.02)	(0.0113)
$\log(tna)$	0.00656***	-0.0163***	-0.00439**	-0.0208***
	(0.0021)	(0.0016)	(0.002)	(0.0035)
age	-1.2e-04*	-2.96e-04***	-0.001***	2.7e-04**
	(6.48e-05)	(4.76e-05)	(7.19e-05)	(1.14e-04)
σ				-23.4***
				(0.845)
Adj. R^2	6.02	7.87	13.46	28.14
Obs.	248227	242426	205665	203646

This table reports estimated coefficients for monthly regressions (41)-(44) between 1980 and 2008. Dependent variable in columns 1-3 is the difference in alpha in two successive months, quarters or years, respectively. In column 4, dependent variable is the difference between sigmas in two successive months. Discussion is in the section 3.1. The coefficients are estimated by pooled OLS regression. Covariance matrix for error terms is a robust estimator in the case of correlated errors. *, ** and *** means significance at 10%, 5% and 1% level, respectively. Standard errors for coefficients are reported in parenthesis. Obs. is the number of observations. All coefficients are in percentage terms.

Table 2.2: Level of fees, alpha and sigma.

Variable	Coefficient in (45)	Coefficient in (45) w/o σ	Coefficient in (46)
σ_{t-1}	0.678***		
	(0.07)		
$\alpha_{t-1,t-12}$	0.131***	0.255***	0.169***
	(0.047)	(0.052)	(0.051)
$\alpha^2_{+,t-1,t-12}$			1.39***
			(0.17)
$\alpha^2_{-,t-1,t-12}$			-0.66***
			(0.11)
$turn_t$	0.011***	0.0126***	0.0124***
	(0.0013)	(0.0014)	(0.0014)
age_{t-1}	1.07e-04***	1.19e-04***	1.16e-04***
	(1.54e-05)	(1.56e-05)	(1.55e-05)
$log(tna_{t-1})$	-0.0104***	-0.0109***	-0.0101***
	(5.78e-04)	(6.04e-04)	(5.98e-04)
Adj. R^2	22.19	19.15	19.96
Obs.	58709	58709	58260

This table reports estimated coefficients for the quarterly/yearly regressions (45) (columns 1-2) and (46) (column 3) between 1980 and 2008. The dependent variable in columns 1, 2 and 3 is the level of fees in a fund. Discussion is in the section 3.2. The coefficients are estimated by pooled OLS regression. Covariance matrix for error terms is a robust estimator in the case of correlated errors. *, ** and *** means significance at 10%, 5% and 1% level, respectively. Standard errors for coefficients are reported in parenthesis. Obs. is the number of observations. All coefficients are in percentage terms.

Table 2.3: Fund flows, alpha and sigma.

Variable	C. in (48)	C. in (49)	C. in (50)	C. in (51)	C. in (48)	C. in (50)
σ	5.69***		7.97***		5.59***	7.88***
	(1.41)		(1.5)		(1.4)	(1.5)
α_{month}	6.24***	6.05***	4.87***	4.91***	6.16***	4.84***
	(0.61)	(0.61)	(0.47)	(0.46)	(0.61)	(0.47)
α_{year}	88.96***	89.59***	85.8***	86.06***	88.87***	85.82***
	(3.35)	(3.54)	(3.84)	(4.16)	(3.36)	(3.84)
$\alpha^2_{+,year}$		6.65		16.63***		
		(4.34)		(4.32)		
$\alpha^2_{-,year}$		-6.86*		-6.64**		
		(3.8)		(3.2)		
$flow_{t-1}$	2.62***	2.42**	5.78***	6.33***	2.6***	5.76***
$(flow_{t-3,t-5})$	(0.98)	(0.93)	(1.72)	(2.17)	(0.97)	(1.72)
$flow_{t-2}$	1.73***	2.27**			1.72***	
	(0.51)	(0.78)			(0.51)	
$fees$	-889.83***	-868.32***	-998.42***	-962.7***	-889.09***	-996.12***
	(50.1)	(49.97)	(54.54)	(54.5)	(50.1)	(54.54)
age	-0.00904***	-0.00868***	-0.00918***	-0.0088***	-0.009***	-0.0092***
	(4.66e-04)	(4.81e-04)	(5.15e-04)	(5.46e-04)	(4.66e-04)	(5.15e-04)
$log(tna)$	0.162***	0.161***	0.11***	0.111***	0.16***	0.111***
	(0.013)	(0.013)	(0.014)	(0.0145)	(0.013)	(0.0141)
$feesdiff_+$					-225.96***	-318.88***
					(75.93)	(70.62)
$feesdiff_-$					-255.96***	-332.52***
					(77.47)	(79.37)
Adj. R^2	13.25	12.55	15.41	14.71	13.25	15.4
Obs.	224388	221313	222269	219112	223837	221734

This table reports estimated coefficients for the monthly regressions (48)-(51) between 1980 and 2008. The dependent variable is monthly fund flows in columns 1, 2 and 5, and average quarterly fund flows in columns 3, 4 and 6. Columns 5 and 6 include fees change as a regressor; discussion is in the sections 3.3 and 3.4. The coefficients are estimated by pooled OLS regression. Covariance matrix for error terms is a robust estimator in the case of correlated errors. *, ** and *** means significance at 10%, 5% and 1% level, respectively. Standard errors for coefficients are reported in parenthesis. Obs. is the number of observations. All coefficients are in percentage terms.

Table 2.4: Difference in fees, alpha and sigma.

Variable	Coefficient in (52), 1981-1999	Coefficient in (52), 2000-2008	Coefficient in (53), 1981-1999	Coefficient in (53), 2000-2008
σ	5.02**	-10.57***	0.26***	-0.269***
	(2.5)	(1.94)	(0.1)	(0.062)
α_{year}	-34.04***	-39.04***	-1.51***	-1.53***
	(4.44)	(3.46)	(0.186)	(0.119)
$fees$	-657.59***	-64.12	-37.06***	-8.49***
	(93.9)	(57.75)	(3.75)	(1.78)
age	0.0288***	0.00313***	1.53e-04***	8.5e-05***
	(0.0009)	(0.0004)	(3.35e-05)	(1.28e-05)
$log(tna)$	-0.227***	-0.122***	-0.00875***	-0.00205***
	(0.03)	(0.178)	(0.00115)	(5.34e-04)
Adj. R^2	5.22	13.81	8.69	17.8
Obs.	3968	9189	3985	9193

This table reports estimated coefficients for the regressions (52) and (53) between 1980 and 2008. The dependent variable in columns 1 and 2 (logit regression) is the sign of difference in fees in a fund at the moment when fees are changed, and in columns 3 and 4 (pooled GLS regression) is the actual change in fees. Discussion is in the section 3.4. Covariance matrix for error terms is a robust estimator in the case of correlated errors. *, ** and *** means significance at 10%, 5% and 1% level, respectively. Standard errors for coefficients are reported in parenthesis. Obs. is the number of observations. Coefficients in columns 1 and 2 are in absolute terms, and coefficients in columns 3 and 4 are in percentage terms.

Table 2.5: Level of fees and change in fees.

Variable	Coefficient in (54), 1981-1999	Coefficient in (54), 2000-2008	Coefficient in (55), 1981-1999	Coefficient in (55), 2000-2008
σ	0.694***	0.666***	0.687***	0.661***
	(0.101)	(0.0777)	(0.101)	(0.0777)
α_{year}	0.158***	0.168***	0.157***	0.153***
	(0.025)	(0.026)	(0.028)	(0.031)
$fees_+$	-0.0054***	0.00779***	-3.49**	11.73***
	(0.00203)	(0.001)	(1.59)	(1.4)
$fees_-$	-0.0076***	-0.0098***	-11.32***	-15.7***
	(0.0019)	(8.82e-04)	(1.48)	(1.09)
age	9.31e-05***	1.17e-04***	9.33e-05***	1.17e-04***
	(2.55e-05)	(1.63e-05)	(2.55e-05)	(1.63e-05)
$log(tna)$	-0.008***	-0.0119***	-0.00803***	-0.00118***
	(8.5e-04)	(6.14e-04)	(8.52e-04)	(6.12e-04)
Adj. R^2	13.55	19.54	13.49	19.56
Obs.	82714	151431	85110	154552

This table reports estimated coefficients for the regressions (54) and (55) between 1980 and 2008. The dependent variable is the level of fees. $fees_+$ and $fees_-$ represent dummy variables in columns 1 and 2, and dummy multiplied by actual fees change in columns 3 and 4. Discussion is in the section 3.4. The coefficients are estimated by pooled OLS regression. Covariance matrix for error terms is a robust estimator in the case of correlated errors. *, ** and *** means significance at 10%, 5% and 1% level, respectively. Standard errors for coefficients are reported in parenthesis. Obs. is the number of observations. All coefficients are in percentage terms.

Table 2.6: Change in fees and future performance.

Variable	Coefficient in (56), 1981-1999, only for changing funds	Coefficient in (56), 2000-2008, only for changing funds	Coefficient in (56), 1981-1999, for funds at the time of change	Coefficient in (56), 2000-2008, for funds at the time of change	Coefficient in (56), 1981-1999, for all funds	Coefficient in (56), 2000-2008, for all funds
$fees_+$	0.853*	1.794***	0.714**	1.736***	0.752**	1.751***
	(0.468)	(0.304)	(0.347)	(0.21)	(0.347)	(0.211)
$fees_-$	0.299***	1.835***	0.575**	1.705***	0.479**	1.728***
	(0.463)	(0.279)	(0.285)	(0.189)	(0.283)	(0.189)
$fees$	0.874**	-0.758***	0.929***	-0.779***	0.395**	-0.805***
	(0.376)	(0.189)	(0.206)	(0.085)	(0.153)	(0.082)
age	5.62e-06	2.87e-06*	1.37e-06	3.77e-06***	4.93e-06***	3.45e-06***
	(4.69e-06)	(1.48e-06)	(2.25e-06)	(7.79e-07)	(1.79e-06)	(6.88e-07)
$log(tna)$	-7.7e-04***	-4.3e-04***	-5.68e-04***	-4.9e-04***	-6.31e-04***	-4.68e-04***
	(1.25e-04)	(5.52e-05)	(6.83e-05)	(3.03e-05)	(5.63e-05)	(2.86e-05)
Adj. R^2	11.33	9.25	7.23	8.57	8.38	7.97
Obs.	3124	8881	11584	59112	65141	148026

This table reports estimated coefficients for the regression (57) between 1980 and 2008. The dependent variable is the change in annual alpha. Columns 1 and 2 consist of funds changing fees. Columns 3 and 4 consist of all fund-month observations at months at which at least one fund changed fees. Columns 5 and 6 consist of all fund-month observations. Discussion is in the section 3.4. The coefficients are estimated by pooled OLS regression. Covariance matrix for error terms is a robust estimator in the case of correlated errors. *, ** and *** means significance at 10%, 5% and 1% level, respectively. Standard errors for coefficients are reported in parenthesis. Obs. is the number of observations. Coefficients are in absolute terms.

3. Human Capital, Training and Portfolio Choice over the Life Cycle

3.1. Introduction

Human capital is the most important non-financial source of income. The wage streams associated with it and housing are the only sources of wealth for the households with no exposure to financial markets. Thus it seems interesting to investigate the behavior of agents who can influence their future incomes through human capital channel. This approach allows to endogenize earnings and consider simultaneous decisions on portfolio, consumption and labor choices.

There are three main ways to alter the human capital of an agent. First and the most important is to get a university degree or an MBA. I will not consider this type of education and will model it with the difference in initial human capital. The second way is "learning-by-doing" which means that a person can change his overall knowledge at work. The last one is to have short training courses, either general or specific to the job or firm in which the agent works. All three ways help to increase future streams of wage income, and they make earnings endogenous on the decision to train or acquire knowledge.

This chapter examines the model in which the agent can choose his time of work and training, as well as consumption and portfolio allocation. With a given probability, he can be unemployed in any period, and he maximizes the expected utility derived from consumption, leisure and final wealth. I assess the ability of the model to replicate general properties of U.S. data on wealth and wage distribution, working time and share of risky asset in the portfolio.

Recent research has shown the ability of life-cycle models to match U.S. data. Yaron et al. (2006) study the model in which agents with learning abilities fixed for the whole life can divide their available time between working for wage and increasing human capital. They set the wage per unit of human capital to be constant and thus endogenize the wage income because of changes in human capital. The paper shows that the cross-sectional distribution of earnings can be matched closely for some initial distribution of abilities and human capital. Yet this paper does not consider consumption/leisure or portfolio choice and does not report the simulated labor supply. Moreover, the authors matched the behavior of the wage income from age 20 to 58 and do not take into account sharp decline in earnings in the ages 60-65 that can not be explained by the model.

Shaw (1989) solves the model of leisure and consumption choice with quadratic function for human capital accumulation and no bequest. The wage is determined from equilibrium conditions. She estimates the parameters of the model using U.S. data and shows that rental rates of human capital should be changing over time to match the labor supply and investment in human capital choice.

Gourinchas and Parker (2002) consider the framework with consumption choice in which they are able to characterize consumption over the life cycle for different groups of consumers taking into account their education. The income process is exogenous and stochastic. The authors show that households save actively in later periods while enjoy consumption on earlier stages of their life. They argue that this behavior is incompatible with benchmark models of representative consumer.

French (2005) builds the model to study labor supply and retirement behavior. He allows the agents to differ in their health and to choose consumption and leisure. The process for wages is stochastic, depends on individual's characteristics and includes autoregressive part. The author shows that wealth accumulation is close to what is observed in the U.S. data, yet the shape of the labor supply in earlier ages can not be reproduced.

Cocco, Gomes and Maenhout (2005) study the models in which wage income is an exogenous yet realistically calibrated process. In these models the agent chooses consumption and portfolio, and wage income plays the role of riskless asset because its return is uncorrelated with the equity return. The findings support common rule of decreasing share of risky asset in the portfolio and generate consumption profiles reasonably close to the U.S. data.

Ameriks and Zeldes (2004) and Wachter and Yogo (2009) provide empirical evidence on portfolio share dependence on wealth and age. It is shown that in Survey of Consumer Finance the share of risky assets in the portfolio rises with wealth and is non-decreasing in age (note, however, that there is a cohort effect that impacts higher share of risky investments for baby-boom generation).

Finally, Gomes, Kotlikoff and Viceira (2008) consider flexible labor supply and portfolio choice model in which leisure and consumption are substitutes and utility is non-separable in these two variables. Authors include retirement in the model, and wage follows exogenous stochastic process. The shapes of labor supply and consumption correspond to the data. Portfolio share behaves as a decreasing function of age until retirement and increases afterwards.

My model is similar to Yaron et al. (2006) in human capital process and differs in three dimensions. First, the agents can use training to raise human capital. This allows the agents to accumulate knowledge faster in younger and middle ages and contributes into the increase in variation in wage income with age. Second, I allow for portfolio and consumption choices. Third, I assume that the agent can be unemployed with i.i.d. rate in any period. This feature of the model influences the resulting share of the risky asset in portfolio and makes it inversely U-shaped rather than decreasing with age.

The main results of this chapter are the following. I show that the levels and shapes of wealth and wage income can be matched. The labor supply is a decreasing function of age and the level of training shows inverse U-shaped behavior. Portfolio holdings exhibit an inverse U-shaped form and are low at earlier years, while consumption is an increasing function of age but is almost flat near retirement. It is also shown that the results are qualitatively robust to changes in parameters.

In the next section I set up the model and shortly discuss the solution tools. In section 3.3 I discuss the data and the parameters for calibration. Section 3.4

summarizes the results, and section 3.5 states some robustness check. Conclusions finalize the chapter in section 3.6.

3.2. The Model

3.2.1. Assets, Wealth and Investment Opportunities

Agents have two types of assets: financial wealth X_t and non-tradable human capital H_t which generates wage income.

There are two financial securities in the economy: a risk-free bond with return R_f and a risky equity with log-normally distributed return R_t. I assume that returns on risky asset are i.i.d. over time periods. The agent decides on the portfolio share of the riskless asset $(1-\alpha_t)$ and α_t is invested in stock.

The agent also receives wage income. I assume that labor market is competitive for workers. It means that the agent is paid the wage W_t for the unit of human capital supplied to the labor market; hence, his income is equal to $Y_t = W_t H_t L_t$. Here L_t is the labor supply. Wage is assumed to follow Markov chain with two states. One of the states is set to be 0 to capture the effect of possible unemployment of the agent and the other depends on time (see section 3.3 for details). I call these two states "unemployment" and "employment", respectively.

Finally, human capital can be increased in two ways. First, the agent experiences "learning-by-doing" and his human capital improves with more time spent on work. Second, he can train his abilities and augment the future expected income stream, yet this is a costly activity both because of a decrease in leisure and direct financial costs (the costs are included into the budget constraint (2)). I denote training time by E_t. The evolution of human capital is assumed to be

$$H_{t+1} = e^{a\beta_H E_t^{\lambda_H}}((1-\delta_t)H_t + aL_t). \tag{71}$$

Here δ_t is the rate of depreciation of human capital and a is the learning ability of the agent; the difference in the latter across agents is shown to be important to explain cross-sectional distribution of wage income in Yaron et al. (2006). As in that paper, I consider a to be fixed for the agent for the whole life cycle.

Compare (71) to what Yaron et al. (2006) use: in their paper the evolution of human capital is

$$H_{t+1} = (1-\delta)H_t + a(H_t(1-L_t))^\eta.$$

The difference is that, first, the "training" or production of human capital $a(H_t(1-L_t))^\eta$ is not a proportional factor which it appears to be in econometric analysis (see discussion of parameters in section 3.3), and the returns on human capital are higher for a given level of labor supply in their model. Second is that time spent on the work is not considered to increase human capital. The final difference that has to be emphasized is that in Yaron et al. (2006) human capital depreciation rate δ is constant over time while in my model it is different in the first and the second half of the life cycle; this allows for a better matching for wage income and wealth accumulation.

Training costs depend on the state of wage: they are equal to c_e for the unit in employment state and c_u for the unit in unemployment state. Denote the costs of training as $c(W_t)$.

The timing of the model is the following. In the beginning of period t the agent knows his wealth X_t, human capital H_t, learning abilities a and wage W_t and chooses portfolio allocation, labor, training and consumption. The agent faces the borrowing constraint and cannot borrow using future income as a collateral, thus $C_t \leq X_t + Y_t$. In the end of the period return on equity is realized.

The budget constraint for financial wealth is then:

$$X_{t+1} = (W_t H_t L_t + X_t - C_t - c(W_t)E_t)(\alpha_t R_t + (1-\alpha_t)R_f). \tag{72}$$

I do not include tax on wage income because this only changes the scale of W_t.

3.2.2. Agents and Preferences

I assume that there is a number J of agents who live for a fixed number of periods T. The agents are different in their learning abilities, initial levels of human capital, financial wealth and investment opportunities. Agents $j = 1, ..., J_1$ can invest in risky equity, while agents $j = J_1 + 1, ..., J$ can not.

Consider first an agent $j = 1, ..., J_1$ who has a given learning ability (a), initial financial wealth ($X_0 > 0$) and human capital ($H_0 \geq 0$), and is able to invest in the risky asset. Each period the agent is endowed with 1 unit of time and maximizes expected utility which is separable in consumption and leisure:

$$\max \mathbf{E}_1 \left[\sum_{s=1}^{T} \beta^s \left(\frac{C_s^{1-\gamma}}{1-\gamma} + \nu \frac{(1 - L_s - E_s)^{1-\lambda_L}}{1-\lambda_L} \right) + \kappa \beta^{T+1} \frac{X_{T+1}^{1-\gamma}}{1-\gamma} \right]. \tag{73}$$

Here C_s is the consumption at time s whereas L_s and E_s are labor supply and training chosen at time s, and I set lower bound for leisure such that $L_s + E_s \leq \overline{L}$. $V(X_{T+1}) = \kappa \frac{X_{T+1}^{1-\gamma}}{1-\gamma}$ captures bequest motives and retirement wealth.

The agent maximizes (73) taking into account (71), (72) and "no default" constraint:

$$X_t \geq 0. \tag{74}$$

Agent $j = J_1 + 1, ..., J$ is unable to invest in the risky asset and he solves the same problem except his share of stock is always 0, $\alpha_t \equiv 0$. For this type of agents randomness is generated by changes in wages. From now on I do not differentiate two types of agents unless it is stated explicitly.

3.2.3. Solution of the Model

The problem can not be solved analytically and is solved numerically using backward induction and Bellman equation.

In the end of the last period the agent has utility $V(X_{T+1})$ and solves the problem in the beginning of period T with constraints (2-4):

$$V_T(X_T, H_T) = \max_{C_T, E_T, L_T, \alpha_T} \left(\frac{C_T^{1-\gamma}}{1-\gamma} + \nu \frac{(1 - L_T - E_T)^{1-\lambda_L}}{1-\lambda_L} + \beta \mathbf{E}_T V(X_{T+1}) \right). \quad (75)$$

The solution is trivial in E_T, since V does not depend on H_{T+1} and thus $E_T = 0$.

In period $s = 1, ..., T-1$ the agent solves the problem with constraints (2-4):

$$V_s(X_s, H_s) = \max_{C_s, E_s, L_s, \alpha_s} \left(\frac{C_s^{1-\gamma}}{1-\gamma} + \nu \frac{(1 - L_s - E_s)^{1-\lambda_L}}{1-\lambda_L} + \beta \mathbf{E}_s V_{s+1}(X_{s+1}, H_{s+1}) \right). \quad (76)$$

The model is solved using grid search, Gaussian quadrature and piecewise shape-preserving cubic interpolation in X_t and linear interpolation in H_t.

3.3. Data and calibration

3.3.1. Data

I use two databases to construct the proxies for wealth, wage income, labor and training. Panel Study of Income Dynamics (PSID) supplies statistics on wealth, labor time and wage income for several years from 1983 to 2004 (I use data for 1983, 1988, 1993, 1998, 2000, 2002 and 2004 due to availability of the data on wealth in these years). The same data was also studied by Yaron et al. (2006). The total number of observations used in the chapter is 8035 for stockholders and 29151 for non-stockholders which is roughly 865 per year. The other database is National Longitudinal Survey of Youth 1979 (NLSY79) that has multiple training questions, namely the length of the training. This data is available for ages 23-49 only. In NLSY79 the number of persons who receive training equals 11148 out of 115543. I set the upper bound for the training to be 400 hours per year and thus the number of observations used to generate first two moments for training is equal to 10024.

I consider the following variables as proxies for the model's variables. There are two wealth proxies I take into account. The first one is the sum of stock and bond holdings. The second one is wide wealth measure including housing, durables and business. I abstract from the cohort effect and consider raw moments of the data. Training proxy is the total time spent on four technical/vocational training courses in NLSY79. This variable is shown to have the most effect on wage increase in Frazis and Loewenstein (2004). I also consider the shape of training reported in the Organization for Economic Co-operation and Development "Economic outlook" (1998), pp.139-140: the share of workers who receive professional and career-upgrading training in the U.S. has a peak at 45-55, slightly decrease afterwards and is increasing from 25 to 45.

Figures 3.1-3.3 show data averages for labor, training, financial wealth (I do not show total wealth), wage income and portfolio share invested in risky asset over the life cycle. Financial and total wealth and wage income are calculated for the respondents of PSID who possess more than 100 and less than 500000 dollars of financial wealth in 1984 real terms. On Figure 3.2 these variables are expressed in U.S. dollars deflated to 2006. Interestingly, the mean of training is almost flat for all the ages available in NLSY79. I put right axis on Figure 3.3 to show number of hours of training per year. The average number of hours spent on training is close to 74 per year over the ages 23-49 and the standard deviation of this mean is approximately 7 (if we exclude two outliers at ages 23 and 48). Labor is restricted to be between 400 and 4000 hours per year. We observe an almost flat labor supply and variable wage and wealth.

Figure 3.4 shows the proportion of agents who invest in risky asset. This share is increasing over the life cycle and averages to 0.3 in PSID. The only variable that can not be explicitly defined for each group of agents is training (because it is not present in PSID). I assume that the fraction of stockholders in the whole population is 0.3.

3.3.2. Parameters: Returns, Utility Function and Discount Rate

Time period in the model is one year. There are $T = 41$ periods, the agent starts his life cycle at age 23 and retires at the beginning of age 66.

The logarithm of the risk-free rate is 1% and risk premium is 4%. Discount rate is $\beta = 0.96$.

The parameters of the utility function for the unconstrained agents in the benchmark model are chosen in such a way that the share of the risky asset is reasonably high and the labor supply is closer to the flat one. I set $\gamma = 8$ and $\lambda_L = 7$ to achieve both goals. Then $\kappa = 20000$ is such that the distribution of wealth is similar to that in the data and $\nu = 3$ is chosen to match average labor supply. We can rewrite the retirement function as $V(X_T) = 43.6\frac{(X/2.4)^{1-\gamma}}{1-\gamma}$ for this parametrization; the average final wealth will be around 2.4 for the unconstrained agents and hence κ is not very high given this scale.

The constrained agents have slightly different utility function, namely I change κ to reflect the fact that consumption streams should be reasonably high and average labor supply is lower. If κ is the same as for the unconstrained agents, the constrained agent will save more to meet the requirement to have higher retirement wealth and will work more. However, in the data non-shareholders work less on average than stockholders. To achieve lower wealth accumulation and lower labor supply, I set $\kappa = 3000$ for the constrained agent.

3.3.3. Parameters: Labor and Training

The (real) wage for employed is set to be equal to $W_T = 0.2$ in the last period (for scaling purpose) and is equal to $W_t = W_t/(1+g)^{T-t}$ where $g = 0.0014$ is chosen to fit PSID growth rate for average real earnings per working person. Since there are two ways to increase human capital and we observe decrease in the real income in the end of the life cycle, the depreciation rate of human capital is set to be higher than in Yaron et al. (2006) and is equal to 4% on average. In the benchmark model I set $\delta_t = 0.02$ for $t = 1, ..., 22$ and $\delta_t = 0.06$ for $t = 23, ..., 41$.

This choice of the parameter δ in the benchmark model makes the decrease in human capital faster in older ages and allows to generate correct shape of wage income. Yet it is disputable, and for the robustness check I solve the model with flat depreciation rate. I show that the shape of earnings can not be properly replicated. In the model with flat rate the peak of wage income is in the fifties rather than in the forties, and this is the result of the proceeding human capital raise in the second part of the life cycle.

There are 90 hours per working week that may be spent on work, training and leisure (we exclude 30 hours the agent may spend on sleep). The lower bound of leisure is set at $\overline{L} = 0.1$ (that is, 27 hours) and thus the agent cannot work and train more than $4500 \times 0.9 = 4050$ hours per year. The upper bound for labor in the model is thus 0.8. Note that average labor supply over the life cycle is around 2125 or 0.48 in PSID. Parameter ν is chosen to match average labor supply of 0.5.

I restrict training time to take no more than 0.098 of total time that adds up to 400 hours per year. This is a reasonable upper bound since in the NLSY79 sample maximum training time is less than 400 for almost 90% of the sample.

I set $\lambda_H = \frac{1}{3}$ as in Frazis and Loewenstein (2004). This paper and several others (Bartel (1995), Booth, Bryan (2005)) have shown that average increase in wage for 60-100 hours of training is $4 - 6\%$; I set $\beta_H = 0.4719$ because higher constant leads to very high levels of training. With this parametrization, 74 hours of training (the average number of training hours in NLSY79) means 6% increase in wages for the agent with the learning ability $a = 0.5$.

The cost of training is more tricky to model because most of the training is paid by the employer (or government if a person is unemployed). Yet there is some evidence (Loewenstein, Spletzer (1998), Booth, Bryan (2005)) that during the training time a person either has lower wage because spends less time on work, no wage if he trains before the job or just shares costs with the employer. Moreover, staying with the same employer reduces the returns on training in comparison to changing the employer after the training. This means that our assumed increase in wages is higher than in reality because there is no change of work in the model. Thus, to generate realistic levels of training I consider explicit cost of training and set it to be $c_e = 1.5$ (or 2500 dollars in 2006 for the training of 74 hours) and $c_u = 0.1$. The parameters are summarized in Table 3.1.

Table 3.1: Parameters of the model

Parameter	**Value**
Discount rate (β)	0.96
Utility function parameter (γ)	8
Utility function parameter (λ_L)	7
Retirement wealth parameter for unconstrained (κ)	43.6
Retirement wealth parameter for constrained (κ)	43.6
Wage increase parameter (β_H)	0.4719
Training power parameter (λ_H)	1/3
Cost of training for employed (c_e)	1.5
Cost of training for unemployed (c_u)	0.1
Leisure utility parameter (ν)	3
Human capital depreciation (δ)	0.02; 0.06
Share of agents able to invest in risky asset	0.30

3.3.4. Calibration

The main goal of the calibration exercise is to match first moments of five variables: wealth X_t, wage income $W_t H_t L_t$, training time E_t, labor time L_t and portfolio share α_t. I concentrate on financial wealth and do not take into account total wealth of the agent that includes housing, business and durables. The total wealth is approximately two times higher than financial wealth in PSID. I report mean and standard deviation of the total wealth W and all the other variables in Table 3.3 for stockholders and in Table 3.7 for bondholders. I summarize the levels for different age groups. It may be better to report financial variables as a ratio to earnings at the same age yet I report them in U.S. dollars deflated to 2006 for simplicity and to make comparison more clear.

I solve the model for a range of possible values of learning ability: $a = 0.3, ..., 0.5$ for unconstrained agents and $a = 0.2, 0.25, ..., 0.4$ for constrained agents. This range of parameters is chosen because of the following arguments. First, the agents with lower abilities would have lower returns on training since the increase in wages for the agent with ability $a = 0.3$ spending 74 hours on training is 3.4% which is low enough. Thus the training will be too expensive for them and they will not train at all. Second, the agent with ability lower than $a = 0.2$ work too much on average, his average working time over the life cycle is close to 0.6. These agents are unable to increase their human capital by means of training, they have low return on learning-by-doing and hence they have to work more to produce reasonable streams of consumption and wealth. Yet this high labor supply contradicts the data.

I simulate the behavior of all the variables. For a given a, initial level of wealth is $X_0 = 0.2$ for agents who can invest in the risky asset (that is, approximately 30000 dollars which corresponds to PSID data) and $X = 0.02$ for agents who can not invest in stock. Also, I start from different levels of human capital that depend on learning ability. For a given a, all the pairs (a, H) have the same probability. I assume that for higher level of learning ability a initial level of human capital is higher on average. This assumption is plausible for two reasons. The agent with higher ability to learn may have better pre-work education. Moreover, the agents with high abilities acquire high levels of human capital during the life cycle. Even if they start with lower initial human capital, they end up with higher human capital and wages. Yet in the data the order is preserved: agents who earn more in the middle and older ages tend to earn more in younger ages. This means that the agents with better learning abilities tend to have higher initial human capital.

Consistent with Figure 3.2, I assume that constrained agents have lower abilities and lower initial human capital on average.

The exact pairs (a, H) are reported in Table 3.2. The choice for the stockholders is motivated by the fact that the initial wage at age 23 is around 30000 dollars, so that around 0.2 in terms of the model and that the maximum wage at ages 38-56 is approximately 2.5 times higher (in U.S. dollars of the year 2006) than initial wage at age 23. Combining these facts, the goal is to choose initial levels of human capital

Table 3.2: Distribution of learning abilities and initial human capital

Value of a	**Values of H, Unconstrained**	**Values of H, Constrained**
0.2	-	0.9, 1.1
0.25	-	0.9, 1.1
0.3	1, 1.2	1, 1.2
0.35	1.2, 1.4	1.2, 1.4
0.4	1.4, 1.6	1.4, 1.6
0.45	1.6, 1.8	-
0.5	1.8, 2	-

such that the behavior of the earnings is inversely U-shaped with a peak at ages 40-49.

The choice for non-stockholders is similar, but I set even lower initial human capital because the wage is lower for this category (25000 dollars at age 23).

3.4. Results for the Benchmark Model

The results for the unconstrained agents are shown in tables 3.4-3.6 and the results for the constrained agents in tables 3.8-3.10. I report average levels of the variables over the life cycle. Labor and training are expressed in fractions of total available time. There is no plausible proxy for consumption in PSID so that I do not report it and do not report standard deviations in the simulations. The standard deviations are expressed in the same units as the variables.

Figures 3.5-3.8 show the average behavior of the respective variables for both types of agents.

Figures 3.5 and 3.6 show the behavior of wage income and wealth for constrained and unconstrained agents, respectively. Financial wealth is accumulated more aggressively in the model than in the data, but this behavior is partially explained by the fact that I do not take into account other types of wealth, for example housing. Total wealth including non-financial wealth is approximately two times higher than financial wealth in PSID and simulated wealth is closer to the total wealth. The overall accumulation of wealth repeats the shape of that in the data until the last period (ages 56-65) when the agents in the model start de-saving. This result is in line with the previous research, for example Gomes, Kotlikoff and Viceira (2008).

Wage income exhibits inverted U-shape as in the data and matches the data closely. The agents increase their human capital and earnings sharply until earlier fiftieth and start to acquire less training and learning after that age. Earnings increase slower in the ages 23-30 than in the data and behave well in the last periods.

The main caveat is consumption. In the utility function, consumption and leisure are compliments; this means that higher leisure leads to higher consumption unlike in Gomes, Kotlikoff and Viceira (2008). This result contradicts the summary in Gourinchas and Parker (2002), yet consumption still shows reasonably flat behavior in the end of the life cycle and does not increase substantially. Also, consumption increases over the life cycle which is consistent with the evidence reported in French (2005) and Gourinchas and Parker (2002).

Figure 3.7 shows the average share of risky security in the portfolio. The model generates too high share of risky asset in the middle ages: while in PSID data the share is close to 0.5 and never reaches 0.6 in the middle ages, simulated proportion of financial wealth in equity is increasing until ages 51-60. Moreover, the share is above 0.7 for the agents with higher learning ability and higher initial level of human capital.

Nevertheless, this result should not evoke much of concern. The inverted U-shape of the risky share in the portfolio is related to the fear of simultaneous unemployment and bad portfolio return which preserves low alpha for the beginning and the end of the life cycle; this is an important result missing in many life-cycle models. At the same time, agents with higher abilities and higher wealth begin to invest more into risky asset because they can cover their consumption in bad times until they become employed and they are eager to collect higher returns on investment. The result is actually close to what is shown in Ameriks, Zeldes (2004) and Wachter, Yogo (2009): the share of risky asset is increasing in wealth and non-decreasing in age. In my model wealth and age are closely related by means of human capital accumulation and thus inverted U-shape pattern is reasonable.

Training and labor are shown in Figure 3.8. Training is not matched closely and follows inversely U-shaped pattern but this is mostly because of the fact that training is not very volatile over the ages in NLSY79 and has approximately the same mean for all ages. Yet the share of agents receiving training increases with age in the model which reflects the ability of the agents to give up a part of consumption and leisure for future earnings. The younger the worker, the less wealth he possesses and the less he is willing to spend time on additional activities except working. This pattern is somehow disputable because Mincer (1989) and Loewenstein, Spletzer (1998) mention that the agents have more training in younger ages than in older ones, but the share of workers who receive training is almost the same as in NLSY79 in all the ages with the mean 8.7% and standard deviation of only 0.6%. Dearden, Reed and Van Reenen (2005) mention that the share of workers receiving training is around 14% in the first half of the 1990th, and Bartel (1995) has a sample with 50% attendance of training. This may mean that the actual share of trained workers and the hours they spend on training is higher than in NLSY79. The study published by Organization for Economic Co-operation and Development (1998) reports the shape similar to the one the model predicts: the training participation peaks for workers at ages 45-54 and slightly decreases for workers at ages 55-65.

Moreover, learning-by-doing involved in the model allows employees to improve their skills even without formal training and thus the number of formally trained agents decreases compared to the model with no learning-by-doing. The overall behavior of the agents is rational: they raise human capital by learning on work until they find it impossible and then start to acquire paid training.

The labor in the model does not behave so close to the data and is too high in the first periods, but qualitatively the results are close to what is shown by French

(2005). The shape of this variable is not replicated, but the overall decrease in labor supply is the same as in PSID.

Second moments of the data are matched to a lesser extent. Wage income and wealth have rising standard deviations until the age 60 but are lower than in the data, and most of the other second moments are also low. Interestingly, the volatilities of labor and training are matched better than the ones of the financial variables. The reason for lower standard deviations in the model than in the data is that the abilities of the agents are very close to each other and thus lead to a small variance in financial variables. Despite that, the overall behavior of variances is like in the data: the difference between the agents increases over time until the retirement age. It would be interesting to see what happens if we allow higher variance in the initial level of human capital and abilities, and this is a step to be performed.

3.5. Robustness Check

To understand what drives the results of the model I implement two robustness checks for the unconstrained agents in the benchmark model. The first one is to set $\alpha = 0.5$ and consider flexible labor supply. The second one is to set $L = 0.5$ (that is, the average labor supply for the whole life cycle) and to derive optimal policy rules under this assumption.

Then I change the utility function to see the dependence of the variables on parameters γ and λ_L. The model is solved for the following parameters of the utility function: $\gamma = 5, \lambda_L = 3$, $\gamma = 8, \lambda_L = 3$, $\gamma = 5, \lambda_L = 7$. The coefficients κ and ν are changed to preserve the same average labor supply for the life cycle and produce similar consumption streams. I do not report them here.

Finally, the model with flat $\delta = 0.04$ is considered. The other parameters are the same as in the benchmark model.

Tables 3.5 and 3.6 contain the optimal solutions for the benchmark model with $\alpha = 0.5$ or $L = 0.5$. We see that the agent who can not change his labor supply is worse off getting lower wage and consumption stream than the unconstrained agent. He accumulates more wealth at the end of the life but has lower utility of leisure. The result comes from lower training and lower share of risky asset in the portfolio: the agent acquires less wealth in younger ages than agent with free labor choice, he is less willing to train in the middle ages and thus end up with the peak of wage at fifties, lower consumption stream and higher wealth at the end of the life cycle.

The agent with constant share of risky asset, however, follows very similar dynamics as one with no constraints, and the reason is that the average share of risky asset in the portfolio over the life cycle is very close to 0.5 in the benchmark model. Labor supply behaves closer to the data though still misses the correct shape (increase in the earlier ages). Wealth is accumulated with almost the same pace until retirement because of fixed share of risky asset.

The results for different parameters of the utility function are shown in tables 3.9-3.11. For both models with $\gamma = 5$, average levels of wealth and wage income seem to be similar to the benchmark model, but the shape of the earnings is skewed to the end of the life cycle, that is, the peak values are produced in the late fifties or early sixties. The share of risky asset, as expected, is higher and close to 1 in the middle ages due to a lower risk aversion, and labor supply is more flat than in the benchmark model when $\lambda_L = 7$ but falls more sharply when $\lambda_L = 3$. Training behaves similarly to the baseline results, and wealth accumulation is more aggressive due to a higher share of risky asset in the portfolio. Consumption follows more reasonable pattern than in the benchmark model and decreases slightly in the end of the life cycle.

For $\gamma = 8$ and $\lambda_L = 3$ results are very close to the benchmark model, but the difference is nonetheless significant. First, the shape of wage income does not follow the data and has a peak in fifties rather than in forties. Second, labor supply declines sharply and matches the data even less.

Finally, I report one model with flat $\delta = 0.4$ for the unconstrained agents, see Table 3.12 (the results for different utility function parameters and constrained agents are qualitatively similar). In this case the peak of earnings is in fifties or sixties and wealth is increasing until the retirement. Thus flat depreciation rate can not produce correct shape of wage income and wealth in the model.

3.6. Conclusions

In this chapter, the model of human capital accumulation and portfolio choice for constrained and unconstrained agents is studied. It is shown that the model can produce qualitatively and quantitatively correct shapes and levels of wage income and wealth. I find that the model generates reasonable patterns for consumption and labor supply. Both share of risky asset in the portfolio and training exhibit inversely U-shaped form. The differences in wages are partially captured by the difference in abilities of the agents.

There are several possible ways to improve the model. First, the model assumes that the wage rate is constant and thus the labor market is competitive. It is interesting to include other types of wage process to reflect business cycle features. The retirement income paid at retirement age can also be included and this may produce more flat behavior of labor supply.

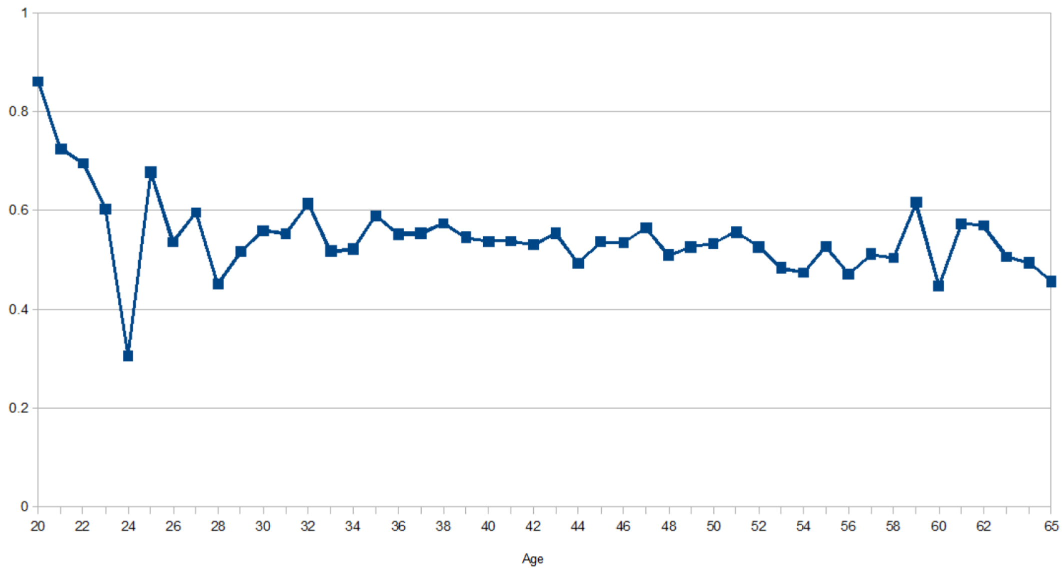
Second, most of the training is explicitly paid by the employer. We may consider the model in which the costs of training are low or even zero. The way to make the agents getting reasonable levels of training hours may be to assign the possibility to train randomly or to assume that the returns to training are random. In this case the agent will choose training time taking into account the loss in leisure and possibility of having no increase in human capital. Final improvement in this direction may be to consider "matching model". Namely, employers may arrange training only to the workers who are productive enough, so that agents with high enough realized abilities could get training. This may also improve training and leisure choice.

Third, we may consider habit-dependent preferences. As is shown in Kiley (2010), preferences with habits allow for a better match for consumption and leisure in U.S. data. This may also make consumption less volatile, and may allow for a reduction in the risk aversion.

Finally, the model may allow for more realistic unemployment rates and include possibility of death or labor inability. This may also change the labor supply and training choices because the agent will take into account the randomness of future earnings.

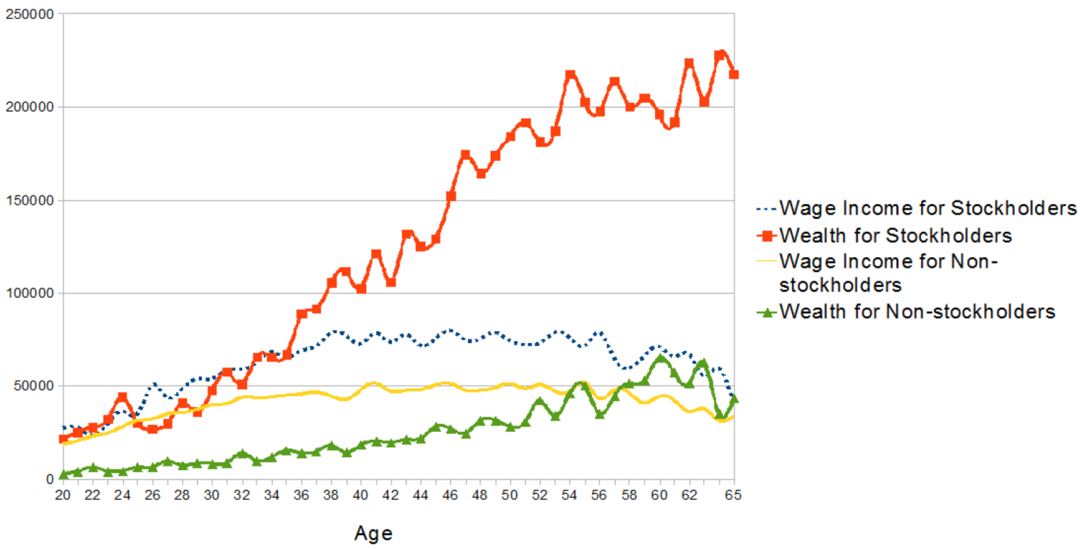
3.7. Figures

Figure 3.1: Share of equity in the portfolio, data.



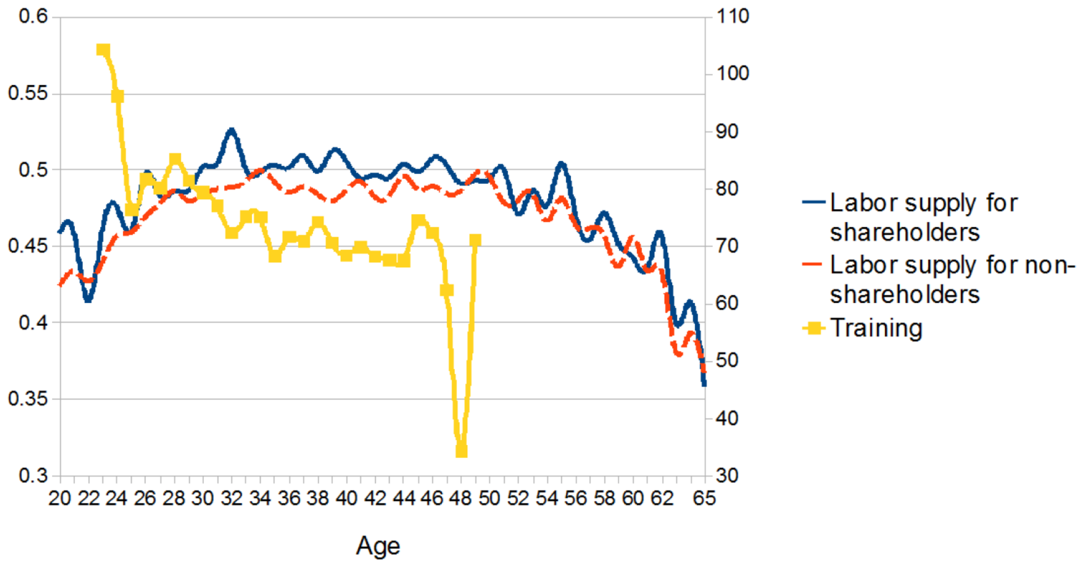
This figure shows average share of equity in the portfolio, depending on age. For the discussion, see section 3.3.

Figure 3.2: Wealth and wage income, data.



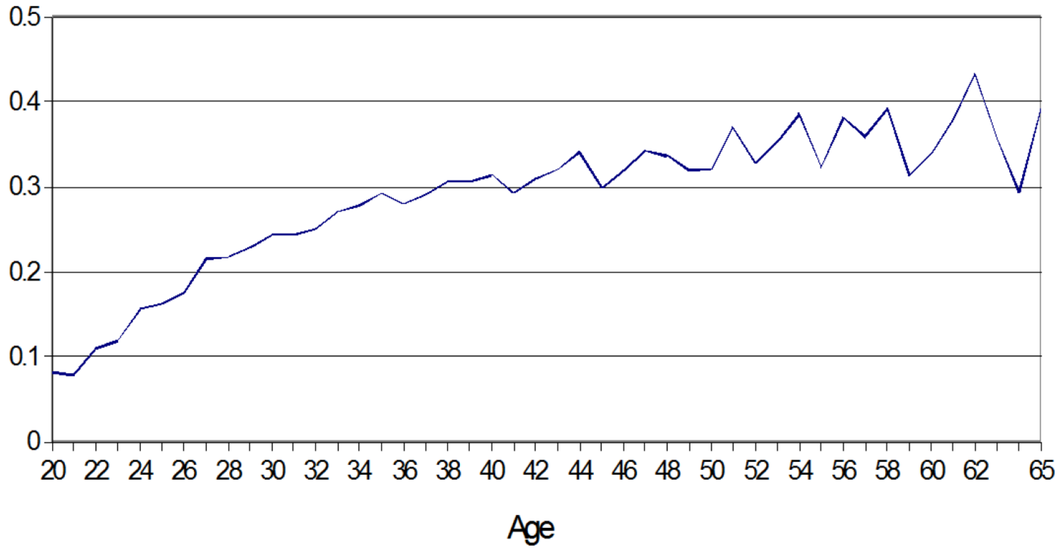
This figure represents average wealth and wage income for stockholders and these who hold bonds only, expressed in U.S. dollars. For the discussion, see section 3.3.

Figure 3.3: Labor supply and training, data.



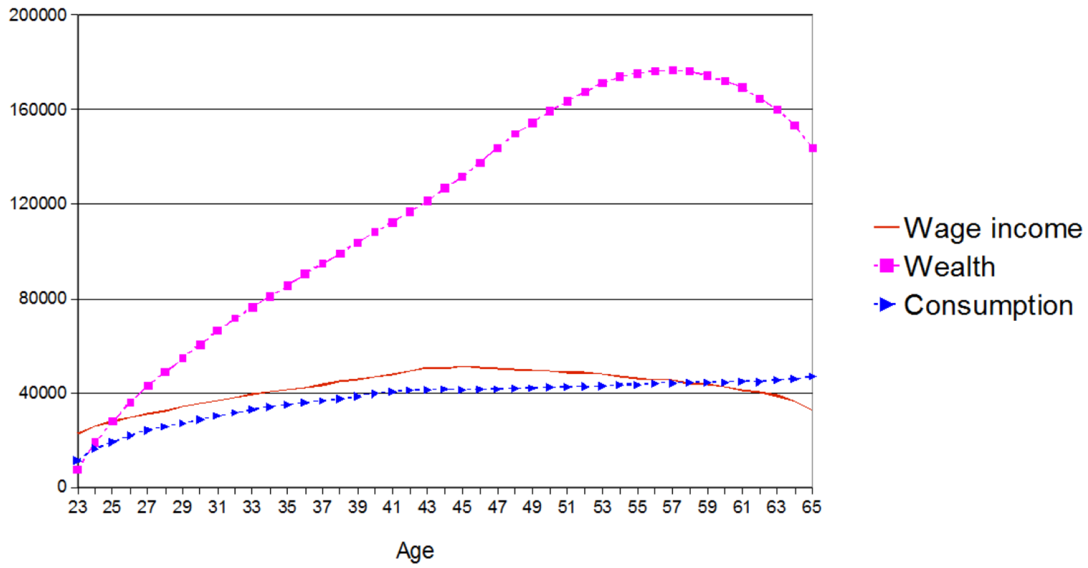
This figure shows average labor supply and training, depending on age. Left axis represents labor supply in fractions of the available time, while right axis represents training in hours per year. For the discussion, see section 3.3.

Figure 3.4: Share of the population investing in stocks, data.



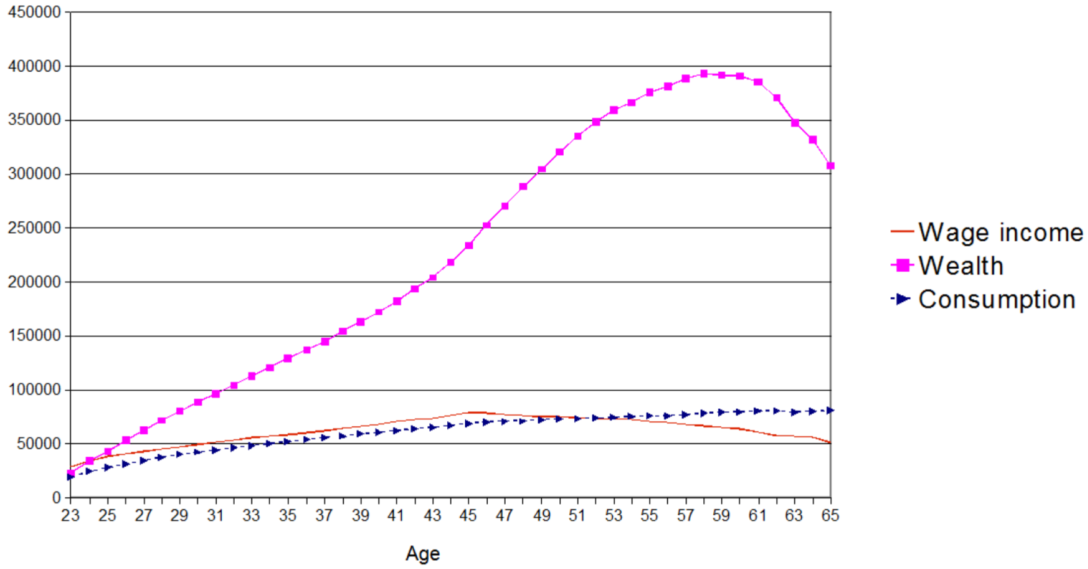
This figure shows the fraction of the population in the U.S. investing in stocks, depending on age. For the discussion, see section 3.3.

Figure 3.5: Wealth, wage income and consumption for constrained agents, simulations.



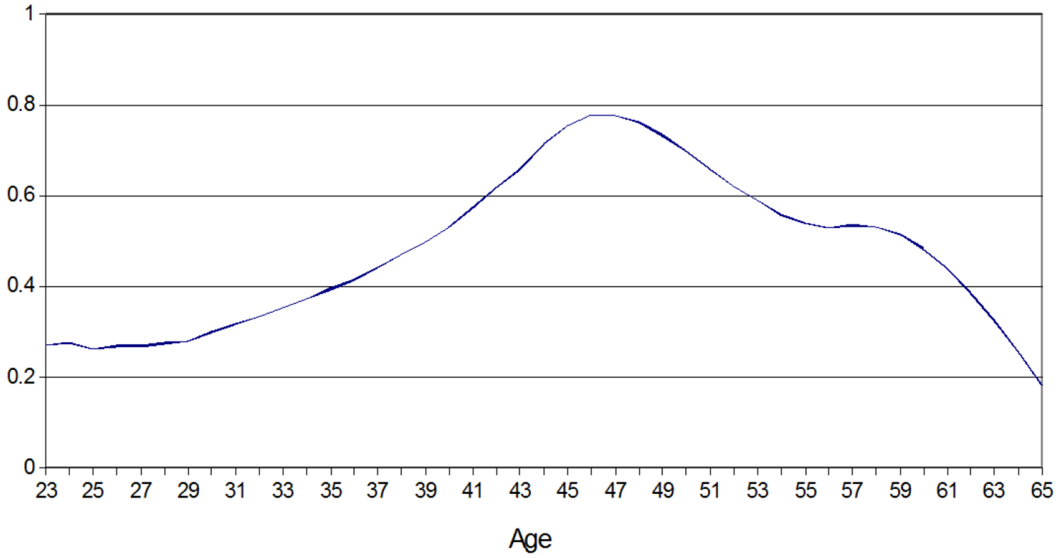
This figure shows average wealth, wage income and consumption for constrained agents, expressed in U.S. dollars, depending on age. For the discussion, see section 3.4.

Figure 3.6: Wealth, wage income and consumption for unconstrained agents, simulations.



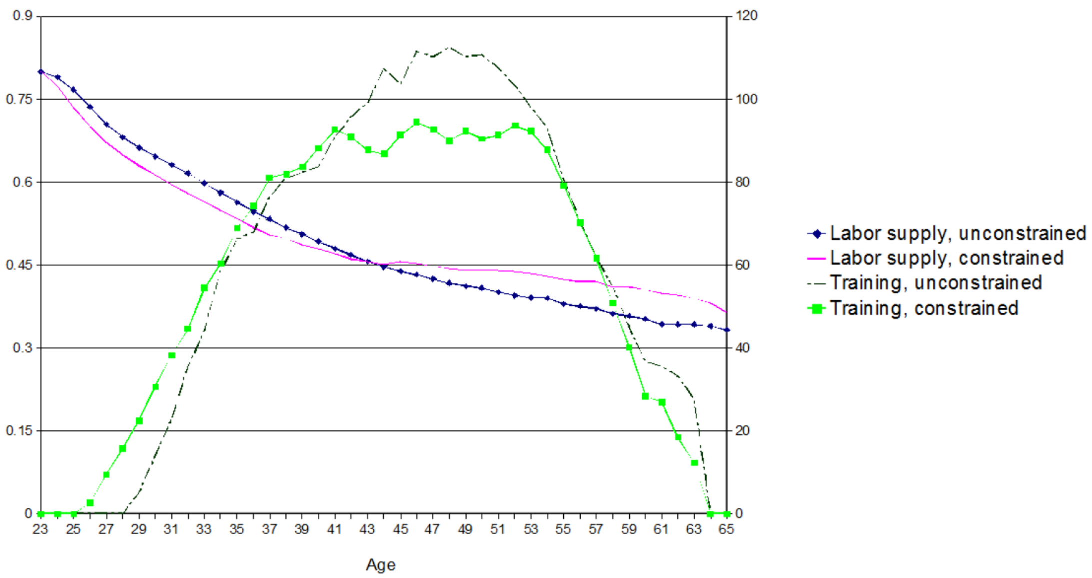
This figure shows average wealth, wage income and consumption for unconstrained agents, expressed in U.S. dollars, depending on age. For the discussion, see section 3.4.

Figure 3.7: Share of risky asset in the portfolio, simulations.



This figure shows share of risky asset in the portfolio for unconstrained agents, depending on age. For the discussion, see section 3.4.

Figure 3.8: Labor and training, simulations.



This figure shows average labor and training for both constrained and unconstrained agents, depending on age. Left axis represents labor supply in fractions of the available time, while right axis represents training in hours per year. For the discussion, see section 3.4.

3.8. Tables

Table 3.3: Data.

	Stockholders, mean						Stockholders, std				
Age	W	X	α	L	E	Y	X	α	L	E	Y
23-30	78173	35980	0.531	0.482	0.016	43940	58779	0.297	0.134	0.006	28666
31-40	199833	80582	0.556	0.506	0.016	68258	133651	0.299	0.116	0.009	51976
41-50	293753	146065	0.532	0.498	0.016	75940	198402	0.304	0.122	0.008	58650
51-60	428738	199037	0.512	0.473	0.016	71189	232081	0.307	0.130	0.009	60486
61-65	567382	212514	0.513	0.418	0.016	57659	254816	0.319	0.145	0.004	49602

This table shows means and standard deviations of the data from PSID and NSLY for stockholders. W is the total wealth including housing, X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply, E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.3.

Table 3.4: Benchmark model.

	Model, mean					Model, std				
Age	X	α	L	E	Y	X	α	L	E	Y
23-30	52669	0.260	0.735	0.00	39794	12656	0.032	0.027	0.00	6091
31-40	125370	0.390	0.574	0.009	58027	37507	0.093	0.052	0.006	9251
41-50	232224	0.690	0.447	0.019	75001	103192	0.17	0.088	0.008	14841
51-60	366383	0.577	0.383	0.017	70889	170805	0.202	0.118	0.005	17411
61-65	355898	0.344	0.342	0.005	57975	138964	0.046	0.148	0.001	17411

This table shows means and standard deviations of simulated data for stockholders. X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply, E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.4.

Table 3.5: Fixed labor supply ($L = 0.5$).

Model, mean						Model, std				
Age	X	α	L	E	Y	X	α	L	E	Y
23-30	32319	0.282	0.5	0	24066	3536	0.064	0	0	4625
31-40	69984	0.300	0.5	0.004	40208	6803	0.037	0	0.003	8825
41-50	161051	0.423	0.5	0.009	65474	15621	0.160	0	0.004	22784
51-60	357865	0.436	0.5	0.007	71022	23955	0.254	0	0.002	27389
61-65	482307	0.286	0.5	0.003	65102	33465	0.142	0	0.002	23190

This table shows means and standard deviations of simulated data for stockholders with fixed labor supply $L = 0.5$. X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply, E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.5.

Table 3.6: Flexible labor supply, fixed share of risky asset ($\alpha = 0.5$).

Model, mean						Model, std				
Age	X	α	L	E	Y	X	α	L	E	Y
23-30	53414	0.5	0.728	0.00	39344	14934	0	0.046	0.00	6656
31-40	129829	0.5	0.572	0.010	57500	40167	0	0.056	0.007	9264
41-50	216497	0.5	0.458	0.021	74166	74924	0	0.082	0.010	15059
51-60	329557	0.5	0.411	0.016	72971	125447	0	0.114	0.007	16922
61-65	344385	0.5	0.355	0.005	59338	120269	0	0.132	0.002	17025

This table shows means and standard deviations of simulated data for stockholders with fixed $\alpha = 0.5$. X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply, E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.5.

Table 3.7: Data.

Non-stockholders, mean						Non-stockholders, std			
Age	W	X	L	E	Y	X	L	E	Y
23-30	41076	6923	0.469	0.016	33034	17368	0.129	0.016	18020
31-40	87437	13910	0.488	0.016	44432	35120	0.121	0.016	29831
41-50	138106	25449	0.489	0.016	49151	58963	0.122	0.016	36821
51-60	198448	45235	0.467	0.016	46799	89984	0.127	0.016	34616
61-65	233138	49971	0.411	0.016	36354	103748	0.144	0.016	28317

This table shows means and standard deviations of the data from PSID and NSLY for non-stockholders. X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply, E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.4.

Table 3.8: Benchmark model.

Model, mean						Model, std				
Age	X	α	L	E	Y	X	α	L	E	Y
23-30	33999	0	0.708	0.001	29245	8532	0	0.032	0.001	4369
31-40	82967	0	0.544	0.008	40873	23244	0	0.058	0.008	6476
41-50	130265	0	0.456	0.015	49714	32457	0	0.078	0.015	8545
51-60	171434	0	0.427	0.010	46681	45255	0	0.088	0.010	9457
61-65	160564	0	0.389	0.003	38818	46643	0	0.106	0.003	6823

This table shows means and standard deviations of simulated data for non-stockholders. X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply, E E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.4.

Table 3.9: $\gamma = 5$, $\lambda_L = 7$.

	Unconstrained agent					Constrained agent			
Age	X	α	L	E	Y	X	L	E	Y
23-30	41655	0.412	0.609	0	39672	25802	0.582	0.001	29011
31-40	86763	0.673	0.523	0.008	59450	57917	0.490	0.008	41996
41-50	180532	0.908	0.437	0.020	81575	95248	0.427	0.015	53276
51-60	328249	0.853	0.402	0.019	85028	137848	0.426	0.012	53713
61-65	433831	0.470	0.391	0.005	79410	182593	0.424	0.004	50483

This table shows means of simulated data for all agents. X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply, E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.5.

Table 3.10: $\gamma = 8$, $\lambda_L = 3$.

	Unconstrained agent					Constrained agent			
Age	X	α	L	E	Y	X	L	E	Y
23-30	54381	0.268	0.777	0.00	42041	34384	0.725	0.002	29162
31-40	140134	0.513	0.589	0.013	60796	78647	0.514	0.014	37836
41-50	267362	0.881	0.416	0.021	71793	135815	0.421	0.021	46385
51-60	367777	0.732	0.320	0.014	58788	165944	0.344	0.015	35928
61-65	286783	0.406	0.289	0.004	44125	117625	0.297	0.003	25697

This table shows means of simulated data for all agents. X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply,E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.5.

Table 3.11: $\gamma = 5$, $\lambda_L = 3$.

	Unconstrained agent					Constrained agent			
Age	X	α	L	E	Y	X	L	E	Y
23-30	40890	0.451	0.724	0.00	39864	36156	0.689	0.002	28922
31-40	95220	0.771	0.575	0.012	59397	82701	0.536	0.014	41423
41-50	202749	0.978	0.445	0.022	78134	142816	0.466	0.020	57013
51-60	347668	0.927	0.358	0.018	71234	174497	0.381	0.017	48921
61-65	318674	0.559	0.340	0.005	59743	123688	0.334	0.003	38466

This table shows means of simulated data for all agents. X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply, E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.5.

Table 3.12: $\gamma = 8$, $\lambda_L = 7$, $\delta = 0.04$, unconstrained agents.

Age	X	α	L	E	Y
23-30	50924	0.257	0.740	0.00	38330
31-40	112272	0.366	0.608	0.007	51984
41-50	185833	0.552	0.507	0.016	63909
51-60	289247	0.516	0.429	0.013	71589
61-65	326900	0.359	0.356	0.005	67087

This table shows means of simulated data for stockholders. X is the financial wealth, α is the fraction of the risky asset in the portfolio, L is the labor supply, E corresponds to training hours, and Y is the wage income. For the discussion, see section 3.5.

References

Ackerman, Frank, Stephen J. DeCanio, Richard B. Howarth, and Kristen Sheeran, 2009, "Limitations of integrated assessment models of climate change", *Climatic Change*, vol. 95, no. 3-4, pp. 297-315.

Adda, Jerome, Christian Dustmann, Costas Meghir, and Jean-Marc Robin, 2006, "Career progression and formal versus on-the-job training", working paper.

Ameriks, John, and Stephen P. Zeldes, "How Do Household Portfolio Shares Vary With Age?", 2004, working paper.

Baker, Erin, Leon Clarke, and Ekundayo Shittu, 2008, "Technical change and the marginal cost of abatement", *Energy Economics*, vol. 30, pp. 2799–2816.

Bali, Turan G., Nusret Cakici, Xuemin (Sterling) Yan, and Zhe Zhang, 2005, "Does Idiosyncratic Risk Really Matter?", *Journal of Finance*, vol. 60, no. 2, pp. 905-929.

Balvers, Ronald J., and Dayong Huang, 2007, "Productivity-based asset pricing: Theory and evidence", *Journal of Financial Economics*, vol. 86, no. 2, pp. 405-445.

Bansal, Ravi, and Amir Yaron, 2004, "Risks for the long run: A potential resolution of asset pricing puzzles", *Journal of Finance*, vol. 59, pp. 1481-1509.

Bansal, Ravi, and Marcelo Ochoa, 2009a, "Temperature, Growth, and Asset Prices", working paper.

Bansal, Ravi, and Marcelo Ochoa, 2009b, "Temperature and Expected Equity Returns", working paper.

Barber, Brad M., Terrance Odean, and Lu Zheng, 2005, "Out of sight, out of mind: The effects of expenses on mutual fund flows", *Journal of Business*, vol. 78, pp. 2095-2120.

Bartel, Ann P., 1995, "Training, Wage Growth and Job performance: Evidence from a Company Database", *Journal of Labor Economics*, vol. 13, no. 3, pp. 401-425.

Belo, Federico, 2010, "Production-Based Measures of Risk for Asset Pricing", *Journal of Monetary Economics*, vol. 57, no. 2.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2007, "Portfolio Choice over the Life-Cycle when the Stock and Labor Markets Are Cointegrated", *Journal of Finance*, vol. LXII, no.5, pp. 2123-2167.

Berk, Jonathan B., and Richard C. Green, 2004, "Mutual Fund Flows and Performance in Rational Markets", *Journal of Political Economy*, vol. 112, no. 6, pp. 1269-1295.

Betz, Regina, and Misato Sato, 2006, "Emissions trading: lessons learnt from the 1st phase of the EU ETS and prospects for the 2nd phase", *Climate Policy*, vol. 6, pp. 351–359.

Bodie, Zvi, Robert C. Merton, and William F. Samuelson, 1992, "Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model", *Journal of Economic Dynamics and Control*, vol. 16, no. 3-4, pp. 427-449.

Boldrin, Michele; Christiano, Lawrence J., and Fisher, Jones D.M., 2001, "Habit Persistence, Asset Returns, and the Business Cycle", *American Economic Review*, vol.91, no. 1, pp. 149-166.

Booth, Alison L., and Mark L. Bryan, 2005, "Testing some Predictions of Human Capital Theory: New Training Evidence from Britain", *The Review of Economics and Statistics*, vol. 85, no. 2, pp. 391-394.

Booth, Alison L., and Mark L. Bryan, 2007, "Who Pays for General Training in Private Sector Britain?", *Research in Labor Economics*, vol. 26, pp. 85-123.

Bosetti, Valentina, Carlo Carraro, Marzio Galeotti, Emanuele Massetti, and Massimo Tavoni, 2006, "WITCH: A World Induced Technical Change Hybrid Model", *The Energy Journal*, Special Issue on Hybrid Modeling of Energy-Environment Policies: Reconciling Bottom-up and Top-down, pp. 13-38.

Bosetti, Valentina, Carlo Carraro, and Massimo Tavoni, 2008, "Delayed Participation of Developing Countries to Climate Agreements: Should Action in the EU and US be Postponed?", working paper.

Buchner, Barbara, and Marzio Galeotti, 2004, "Climate policy and economic growth in developing countries: the impact of Kyoto", *International Journal of Global Environmental Issues*, vol. 4, no. 1/2/3, pp. 109-138.

Buchner, Barbara, Carlo Carraro, and A Denny Ellerman, 2006, "The Allocation of European Union Allowances. Lessons, Unifying Themes and General Principles", *CEPR Discussion Papers*, no. 5843.

Bushnell, James B., Howard Chong, and Erin T. Mansur, 2009, "Profiting from regulation: An event study of the EU carbon market", NBER working paper 15572.

Busse, Jeffrey A., and Paul J. Irvine, 2006, "Bayesian Alphas and Mutual Fund Persistence", *Journal of Finance*, vol. 61, no. 5, pp. 2251-2288.

Campanalea, Claudio, Rui Castrob, and Gian Luca Clementi, 2010, "Asset pricing in a production economy with Chew–Dekel preferences", *Review of Economic Dynamics*, vol. 13, no. 2, pp. 379-402.

Carey, David, 2010, "Implementing cost-effective policies in the United States to mitigate climate change", working paper.

Carhart, Mark M., 1997, "On persistence in mutual fund performance", *Journal of Finance*, vol. 52, no. 1, pp. 57-82.

Cocco, Joao F., Francisco J. Gomes, and Pascal J. Maenhout, 2005, "Consumption and Portfolio Choice over the Life Cycle", *Review of Economic Studies*, vol. 18, no. 2, pp. 491-533.

Cooley, T., Prescott, E., 1995. "Economic growth and business cycles", in: Cooley, T. (Ed.), "Frontiers of Business Cycle Research", *Princeton Univ. Press*, pp. 1–38.

Coria, Jessica, 2009, "Taxes, permits, and the diffusion of a new technology", *Resource and Energy Economics*, vol. 31, no. 4, pp. 249-271.

Cremers, Martijn, and Antti Petajisto, 2009, "How Active Is Your Fund Manager? A New Measure That Predicts Performance", *Review of Financial Studies*, vol. 22, no. 9, pp. 3329-3365.

Cuoco, Domenico, and Ron Kaniel, 2006, "Equilibrium Prices in the Presence of Delegated Portfolio Management", working paper.

Dearden, Lorraine, Howard Reed, and John Van Reeven, 2005, "The Impact of Training on Productivity and Wages: Evidence from British Panel Data", working paper.

Den Haan, Wouter J., and Albert Marcet, 1994, "Accuracy in Simulations", *The Review of Economic Studies*, vol. 61, no. 1, pp. 3-17.

Dimson, Elroy, Stefan Nagel, and Garrett Quigley, 2003, "Capturing the Value Premium in the United Kingdom", *Financial Analysts Journal*, vol. 59, no. 6, pp. 35-45.

Energy Information Administration (EIA), 2005, "Documentation for Emissions of Greenhouse Gases in the United States".

Fama, Eugene F., and Kenneth R. French, 2002, "The Equity Premium", *Journal of Finance*, vol. 57, pp. 637-659.

Fama, Eugene F., and Kenneth R. French, 2008, "Mutual Fund Performance", working paper.

Faraglia, Elisa, Albert Marcet, and Andrew Scott, 2010, "Dealing with Maturity: Implications for Debt Management and Fiscal Policy", working paper.

Ferson, Wayne E., and Schadt, Rudi W., 1996, "Measuring fund strategy and performance in changing economic conditions", *Journal of Finance*, vol. 51, pp. 425-461.

Fischer, Carolyn, and Richard G. Newell, 2008, "Environmental and technology policies for climate mitigation", *Journal of Environmental Economics and Management*, vol. 55, pp. 142–162.

Fowlie, Meredith, 2009, "Allocating Emissions Permits in Cap-and-trade Programs: Theory and Evidence", working paper.

Fowlie, Meredith, 2010, "Emissions Trading, Electricity Restructuring, and Investment in Pollution Abatement", *American Economic Review*, vol. 100, no. 3, pp. 837–869.

Frazis, Harley and Mark A. Loewenstein, 2005, " Reexamining the Returns to Training: Functional Form, Magnitude, and Interpretation", *The Journal of Human Resources*, vol. 40, no. 2, pp. 453-476.

French, Eric, 2005, "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour", *Review of Economic Studies*, vol. 72, no. 2, pp. 395-427.

French, Kenneth R., 2008, "The Cost of Active Investing", working paper.

Gaspar, Jose-Miguel, Massimo Massa, and Pedro Matos, 2006, "Favoritism in mutual fund families? Evidence on strategic cross-fund subsidization", *Journal of Finance*, vol. 61, no. 1, pp. 73-104.

Gil-Bazo, Javier, and Pablo Ruiz-Verdu, 2009, "The Relation between Price and Performance in the Mutual Fund Industry", *Journal of Finance*, vol. 64, no. 5, pp. 2153-2183.

Goyal, Amit, and Pedro Santa-Clara, 2003, "Idiosyncratic risk matters!", *Journal of Finance*, vol. 58, pp. 975-1008.

Gomes, Francisco J., and Alexander Michaelides, 2005, "Optimal Life-Cycle Asset Allocation: understanding the Empirical Evidence", *Journal of Finance*, vol. 60, pp. 869-904.

Gomes, Francisco J., and Alexander Michaelides, 2008, "Asset Pricing with Limited Risk Sharing and Heterogeneous Agents", *The Review of Financial Studies*, vol. 21, no. 1, pp. 415-448.

Gomes, Francisco J., Laurence J. Kotlikoff and Luis M. Viceira, 2008, "Optimal Life-Cycle Investing with Flexible Labor Supply: A Welfare Analysis of Life-Cycle Funds", *American Economic Review: Papers and Proceedings*, vol. 98, pp. 297-303.

Gourinchas, Pierre-Oliver, and Jonathan A. Parker, 2002, "Consumption over the Life Cycle", *Econometrica*, vol. 70, no. 1, pp. 47-89.

Gregory, Alan, Rajesh Tharyan, and Angela Christidis, 2009, "The Fama-French and Momentum Portfolios and Factors in the UK", working paper, University of Exeter Business School.

Grishchenko, Olesya V., 2010, "Internal vs. external habit formation: The relative importance for asset pricing", *Journal of Economics and Business*, vol. 62, pp. 176–194.

Gruber, Martin J., 1996, "Another puzzle: The growth in actively managed mutual funds", *Journal of Finance*, vol. 51, no. 3, pp. 783-810.

Hadi, Ali S., 1994, "A modification of a method for the detection of outliers in multivariate samples", *Journal of the Royal Statistical Society, Series B-Methodological*, vol. 56, pp. 393-396.

Hart, Rob, 2008, "The timing of taxes on CO_2 emissions when technological change is endogenous", *Journal of Environmental Economics and Management*, vol. 55, pp. 194–212.

Heinkel, Robert, Alan Kraus, and Josef Zechner, 2001, "The effect of green investment on corporate behavior", *Journal of Financial and Quantitative Analysis*, vol. 35, pp. 431-449.

Heller, Thomas C., and P. R. Shukla, 2003, "Development and Climate: Engaging Developing Countries" – Chapter in *"Beyond Kyoto: Advancing the International Effort Against Climate Change*, J. E. Aldy et al., Eds. (Pew Center on Global Climate Change, Arlington, VA).

Heyerdahl-Larsen, Christian, 2010, "Asset Prices and Real Exchange Rates with Deep Habits", working paper.

Hong, Harrison, and Marcin Kacperczyk, 2007, "The Price of Sin: The Effects of Social Norms on Markets", working paper.

Horowitz, John K., 2009, "The Income-Temperature Relationship in a Cross-Section of Countries and its Implications for Predicting the Effects of Global Warming", *Environmental and resource economics*, vol. 44, no. 4, pp. 475-493.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang, 2010, "Risk Shifting and Mutual Fund Performance", working paper.

Hugonnier, Julien, and Ron Kaniel, 2008, "Mutual Fund Portfolio Choice in the Presence of Dynamic Flows", working paper.

Investment Company Institute, 2009, "Profile of mutual fund shareholders".

Jaffe, Judson, and Robert N. Stavins, 2008, "Linkage of Tradable Permit Systems in International Climate Policy Architecture", *NBER working paper*, no. 14432.

Jermann, Urban J., 1998, "Asset pricing in production economies," *Journal of Monetary Economics*, vol. 41, no. 2, pp. 257-275.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, "On the Industry Concentration of Actively Managed Equity Mutual Funds", *Journal of Finance*, vol. 60, no.4, pp. 1983-2011.

Kiley, Michael T., 2010, "Habit Persistence, Non-separability Between Consumption and Leisure, or Rule-of-thumb Consumers: Which Accounts for the Predictability of Consumption Growth?", *The Review of Economics and Statistics*, vol. 92, no. 3, pp. 679–683.

Klepper, Gernot, and Sonja Peterson, 2004, "The EU Emissions Trading Scheme Allowance Prices, Trade Flows and Competitiveness Effects", *European Environment*, vol. 14, no. 3, pp. 201-218.

Kort, P. M., 1994, "Effects of Pollution Restrictions on Dynamic Investment Policy of a Firm", *Journal of Optimization Theory and Applications*, vol. 83, no. 3, pp. 489-509.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White, 2006, "Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis", vol. 61, no. 6, pp. 2551-2595.

Kaltenbrunner, Georg, and Lars A. Lochstoer, 2010, "Long-Run Risk through Consumption Smoothing", *Review of Financial Studies*, vol. 23, no. 8, pp. 3190-3224.

Loewenstein, Mark A., and James R. Spletzer, 1998, "Dividing the Costs and Returns to general Training", *Journal of Labor Economics*, vol. 16, no.1, pp. 142-171.

Lou, Dong, 2010a, "A Flow-Based Explanation for Return Predictability", working paper.

Lou, Dong, 2010b, "A Test of the Self-Serving Attribution Bias: Evidence from Mutual Funds", working paper.

Martin, Ralf, Laure B. de Preux, and Ulrich J. Wagner, 2009, "The impact of the Climate Change Levy on Business: Evidence from microdata", working paper.

McKibbin, Warwick J., Martin Ross, Robert Shackleton, and Peter J. Wilcoxen, 1999, "Emissions Trading, Capital Flows and the Kyoto Protocol", *Brookings Discussion Papers in International Economics*, no. 144.

Mercer, 2011, "Climate Change Scenarios - Implications for Strategic Asset Allocation", public report.

Milliman, Scott R., and Raymond Prince, 1989, "Firm Incentives to Promote Technological Change in Pollution Control", *Journal of Environmental Economics and Management*, vol. 17, pp. 247-265.

Mincer, Jacob, 1988, "Job Training, Wage Growth, and Labor Turnover", *NBER Working Paper*, no. 2690.

Mincer, Jacob, 1989, "Job Training: Costs, Returns, and Wage Profiles", *NBER Working Paper*, no. 3208.

Neuhoff, Karsten, Kim Keats Martinez, and Misato Sato, 2006, "Allocation, incentives and distortions: the impact of EU ETS emissions allowance allocations to the electricity sector", *Climate Policy*, vol. 6, pp. 73–91.

Newbold, Stephen C., and Adam Daigneault, 2009, "Climate Response Uncertainty and the Benefits of Greenhouse Gas Emissions Reductions", *Environmental and Resource Economics*, vol. 44, no. 3, pp. 351-377.

Newell, Richard G., Adam B. Jaffe, and Robert N. Stavins, 2006, "The effects of economic and policy incentives on carbon mitigation technologies", *Energy Economics*, vol. 28, pp. 563–578.

Oberndorfer, Ulrich, 2009, "EU emission allowances and the stock market: Evidence from the electricity industry", *Ecological Economics*, vol. 68, pp. 1116-1126.

Organization for Economic Co-operation and Development, OECD, 1998, "Employment Outlook", *OECD Publishing*

Organization for Economic Co-operation and Development, OECD, 2009, “The Economics of Climate Change Mitigation: Policies and Options for Global Action Beyond 2012”, Paris.

Pavlova, Anna, and Roberto Rigobon, 2007, "Asset Prices and Exchange Rates", *Review of Financial Studies*, vol. 20, pp. 1139-1181.

Pavlova, Anna, and Roberto Rigobon, 2008, "The Role of Portfolio Constraints in the International Propagation of Shocks", *Review of Economic Studies*, vol. 75, pp. 1215-1256.

Pollet, Joshua M., and Mungo Wilson, 2008, "How Does Size Affect Mutual Fund Behavior?", *Journal of Finance*, vol. 63, no. 6, pp. 2941-2969.

Rabl, Ari, and Bob van der Zwaan, 2009, "Cost–benefit analysis of climate change dynamics: uncertainties and the value of information", *Climatic Change*, vol. 96, no. 3, pp. 313-333.

Ravn, Morten, Stephanie Schmitt-Grohe, and Martin Uribe, 2006, "Deep Habits", *Review of Economic Studies*, vol. 73, pp. 195-218.

Requate, Till, 2005, "Dynamic incentives by environmental policy instruments – a survey", *Ecological Economics*, vol. 54, pp. 175–195.

Reuter, Jonathan, and Eric Zitzewitz, 2010, "How Much Does Size Erode Mutual Fund Performance? A Regression Discontinuity Approach", *NBER working paper*, no. 16329.

Rubin, Jonathan D., 1996, "A Model of Intertemporal Emission Trading, Banking, and Borrowing", *Journal of Environmental Economics and Management*, vol. 31, pp. 269-286.

Shaw, Kathryn L., 1989, "Life-Cycle Labor Supply with Human Capital Accumulation", *International Economic Review*, vol. 30, no. 2, pp. 431-456.

Sirri, Erik R., and Peter Tufano, 1998, "Costly Search and Mutual Fund Flows", *Journal of Finance*, vol. LIII, no. 5, pp. 1589-1622.

Vayanos, Dimitri, and Paul Woolley, 2008, "An Institutional Theory of Momentum and Reversal", working paper.

Veith, Stefan, Joerg Werner, and Jochen Zimmermann, 2009, "Economic Consequences of Emission Trading Schemes: Evidence from the European Power Sector," *Energy Economics*, vol. 31, no.4, pp. 605-613.

Vissing-Jorgensen, Annette, and Orazio P. Attanasio, 2003, "Stock-Market Participation, Intertemporal Substitution, and Risk-Aversion", *American Economic Review*, vol. 93, no. 2, pp. 383-391.

Wachter, Jessica A., and Motohiro Yogo, 2009, "Why Do Household Portfolio Shares Rise in Wealth?", working paper.

Wahal, Sunil, and Albert (Yan) Wang, 2011, "Competition Among Mutual Funds", *Journal of Financial Economics*, vol. 99, no. 1, pp. 40-59.

Weitzman, Martin L., 2009a, "On Modeling and Interpreting the Economics of Catastrophic Climate Change", *The Review of Economics and Statistics*, vol. 91, no. 1, pp. 1-19.

Weitzman, Martin L., 2009b, "Some Basic Economics of Climate Change", working paper

Weitzman, Martin L., 2009c, "The Extreme Uncertainty of Extreme Climate Change: An Overview and Some Implications", working paper

Wermers, Russ, 2000, "Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses", *Journal of Finance*, vol. 55, no. 4, pp. 1655-1703.

Yaron, Amir, Mark Huggett, and Gustavo Ventura, 2006, "Human capital and earnings distribution dynamics", *Journal of Monetary Economics*, vol. 53, no. 2, pp. 265-290.