UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on Macroeconomics and Information Economics

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Economics

by

Jorge Andres Zambrano Riveros

2012

Abstract of the Dissertation

Essays on Macroeconomics and Information Economics

by

Jorge Andres Zambrano Riveros

Doctor of Philosophy in Economics

University of California, Los Angeles, 2012

Professor Andrew Granger Atkeson, Chair

My research interests are in the intersection of macroeconomics and information economics. I am particularly interested in the optimal design of incentives for individuals who interact over time in uncertain environments with imperfect information, and its consequences for aggregates. The dissertation is composed by three chapters that address three particular environments within this area.

The first chapter studies the optimal contract in a principal-agent model where a risk-neutral principal delegates to a risk-neutral agent the decision of whether to pursue a risky project or a safe one. The return from the risky project is unknown and the agent can acquire costly unobservable information about it before taking the decision. The optimal contract suggests that the principal should only reward the agent for outcomes that are significantly better than the safe return. It is also optimal to distort the project choice in favor of the risky one as a mechanism to induce the direct revelation of the uncertain state. In a managerial context, the findings explain why options and profit sharing compensation induce better decision making from CEOs, as well as why excessive risk taking might be optimal.

The second chapter explores the role of effort and human capital as mechanisms to alleviate the idiosyncratic risk faced by individuals in the presence of incomplete markets. I construct a DSGE model where effort and human capital determine the probability of being

employed the next period. I show how in the stationary equilibrium individuals diversify between these mechanisms. As a result, I obtain a wealth distribution that better approximates the real one. The results shed light on the potential implication of combining policies of unemployment insurance and subsidies to education to improve the wealth distribution.

The third chapter studies dynamic stochastic models where current actions are constrained by a current state and determine the distribution over future states. The purpose of this paper is to provide a general learning process that allows agents to take the optimal decision when these endogenous transitions are unknown. Our paper generalizes previous results by not imposing parametric restrictions on the unknown transition functions.

The dissertation of Jorge Andres Zambrano Riveros is approved.

Sushil Bikhchandani

Rosa Liliana Matzkin

Hugo Andres Hopenhayn

Andrew Granger Atkeson, Committee Chair

University of California, Los Angeles

2012

To my family, my friends and my lovely companion. To my omnipresent family, who provided me all their love and guidance every single moment of my life. To my motivating friends, who not only have encouraged me to be curious, but have always joined me in such adventure. To my beautiful Ana Maria, who has been my main support during all these years, who day by day has shown me there is no limit for love.

Table of Contents

List of Figures

Acknowledgments

The author is indebted to Andrew Atkeson, Hugo Hopenhayn and Rosa Matzkin for their helpful advice during all these years. I also gratefully acknowledge the comments from Sushil Bikhchandani, Simon Board, Francisco Buera, Moshe Buchinsky, Maria Casanova, Jernej Copic, David Bardey, Roger Farmer, Jinyong Hahn, Christian Hellwig, Fernando Leibovici, Gonzalo Llosa, Jose Ignacio Lopez, Maurizio Mazzocco, Moritz Meyer-ter-Vehn, Ichiro Obara, Lee Ohanian, Joe Ostroy, Marek Pycia, John Riley, Connan Snider, Venky Venkateswaran, Pierre-Olivier Weill, and Bill Zame, as well as participants of the student seminars at UCLA, and presentations at Universidad del Rosario, Universidad de los Andes, Universidad Carlos III, Universitat Pompeu Fabra, the Midwest Theory Conference, the Midwest Macro Conference, the Midwest Economics Association Conference, Colombian Central Bank and The Guanajuato Workshop for Young Economists. Funding from the Colombian Central Bank and NBER is also acknowledged. Remaining errors are mine.

Vita

Degrees

2009	M.A. Economics, UCLA
2005	M.A. Economics, Universidad del Rosario, Colombia
2003	B.A. Economics, Universidad del Rosario, Colombia

Employment History

2010-2011	Instructor for Game Theory; UCLA
2008-2011	T.A. for Graduate General Equilibrium, and Game Theory; UCLA T.A for Undergrad Micro, Intermediate Micro, and Statistics; UCLA
2009-2010	Research Assistant for Prof. Rosa Matzkin, UCLA
2008	Research Assistant for Prof. John Riley, UCLA
2004-2007	Acting Assistant Professor, Department of Economics, Universidad del Rosario
2004	Junior Researcher, Department of Economics, Universidad del Rosario
2002-2004	Research Assistant, Department of Economics, Universidad del Rosario
2003	Personal Assistant of the Direction, COLCIENCIAS

2003 Research Assistant, ASSALUD

Publications

Zambrano A, Ramirez M, Yepes F, Guerra J, Rivera D. (2008) "What do Living Standards Surveys Show about the Health System in Colombia." *Public Health Reports* 24 (1): 122-130.

Castano RA, Zambrano A. (2006). "Biased Selection within the Social Health Insurance Market in Colombia". *Health Policy* 79(3): 313-324.

Arguello R, Zambrano A. (2006) Is there a poverty trap in the rural sector in Colombia? *Desarrollo y Sociedad* 58: 85-114

CHAPTER 1

Motivating Information Acquisition and (Almost) Good Decisions

1.1 Introduction

It has been argued that much of the financial crisis has been associated to excessive risk taking from CEOs and that such risk taking is misaligned with shareholders interests. Moreover, it has been suggested that option-type contracts are the cause of this misalignment ([21]). This paper uses a model of delegated expertise to explain why stock options and profit sharing are optimal forms of compensation when a CEO has to be motivated to take good decisions. It also provides an explanation of why excessive risk taking is optimal from a shareholder perspective, where excessive risk taking is understood here as pursuing a risky project even though it is ex-ante inefficient.

Models of delegated expertise were first proposed by [41] and [20]. In such models a principal must hire an agent to decide between a risky and a safe project. Before taking the decision, the agent can acquire unobservable information about the risky project by exerting unobservable effort. This framework shares features of moral hazard and hidden information, and incentives must be used to motivate both information acquisition and the (partial) revelation of the obtained information.

Incentives in this scenario are potentially different than the ones in standard moral hazard problems since effort does not generate greater expected returns directly. In contrast, the unobservable action taken by the agent generates a privately observed signal that improves the decision-making. The studied setup is very general and the only restriction imposed is that signals can be ordered in the likelihood ratio order ([53]). Under limited liability for both individuals, the optimal wages suggest that the agent should be rewarded with the return of the risky project only if it is significantly better than the safe return. Moreover, it is optimal to distort the project choice in favor of the risky project as a strategy to reveal the uncertainty directly.

The intuition why contracts reward experimenters only for extreme good outcomes has two elements whose main instrument is the probability of adopting the risky project given the observed return. The first element is a likelihood ratio that unravels the moral hazard concerns. It suggests that we should reward the agent when the probability of choosing the risky project was greater when effort was exerted rather than when no information was acquired. Since such probability is increasing in the risky return when effort was exerted and constant when that was not the case, the moral hazard incentives must be monotone increasing.

There is a second component associated with the adverse selection problem and is summarized by a hazard rate. The principal must also provide incentives for the agent to choose the risky project whenever he observes a sufficiently high signal. Since signals are ordered and wages are monotone, there will exist a unique cutoff signal where an agent is indifferent between the safe and the risky project. Higher signals will induce the choice of the risky project. This implies that, in the limit, the principal wants to penalize agents that chose the risky project when the cutoff signal was observed. Given the ordering of the signals, the probability of being at the cutoff signal given that the risky project was chosen (the hazard rate) is decreasing in the risky return. Therefore, this effect also suggests that compensation should be increasing in risky return.

The intuition for the project choice distortion relies on the idea that the principal has two mechanisms to induce the revelation of the unknown return. The first one is through risky wages as discussed before. The second one is through the safe wage, which is equivalent to picking the cutoff signal. Instead of paying the agent more to reveal the uncertainty imperfectly (the signal), the principal can also decrease the safe wage to induce the choice of the risky project and have the uncertainty revealed directly.

The model applies to a variety of situations. The one principal one agent problem is tightly connected to the optimal compensation of CEOs who must be motivated by the shareholders to undertake risky projects that could potentially lead to higher returns. In this scenario, the effort exerted by the CEO in learning about the portfolio of projects and the learned information is usually never observed by the shareholders, only the project chosen and the realized returns are observed. In this environment, the optimal contract can be implemented using restricted stocks conditional on a performance threshold. When the optimal contract is constrained to be monotone as in [37], the optimal monotone contract is an option with strike price greater than the return of the safe project.

The conclusions derived from the model explain why options and restricted stock are so widely used in this context. It also suggests that it is the best form of compensation to align the interests of shareholders and CEOs. Moreover, the optimal distortion on the project choice suggests that it is on the shareholders interest to have a CEO taking excessive risks. Letting a CEO pursue risky projects that are not ex-ante profitable would help the shareholders ameliorate the information problems.

The setup is also similar to the problem faced by managers who must encourage innovation among her employees to increase the profits of the firm. Workers have to divide their time between undertaking known tasks or exploring new ideas. However, innovations are risky ventures with a high probability of failure, thus agents prefer to put more effort on known tasks which returns are well known. Again issuing options and restricted stock to workers are shown to be useful to encourage more risk-taking.

The problem also resembles the technology adoption decision faced by farmers in developing countries. [26] studied the adoption of HYVs in India and found that imperfect knowledge about the management of the new seeds was a significant barrier to adoption. Hence, the rate of adoption was much slower than the desired one. In this context the amount of trials performed by farmers as well as the information gathered by them are usually unobservable to the social planner. Pricing policies resembling the optimal contract would encourage the adoption of new technologies.

As a last example, the proposed framework also fits the situation of industries with high levels of innovation such as pharmaceuticals. Pharmaceuticals must invest in potential drugs with unknown effectiveness. In the absence of property rights, free-riding reduces innovation and the potential discovery of new drugs. This environment is close to the one studied in this paper since the investment chosen by firms in the earlier stage of research is usually not disclosed, nor are the results from such investments. Given these constraints, the optimal contract suggests the use of patents to motivate innovation. However, such patents must be issued only for breakthroughs and not marginal innovations.

1.1.1 Literature Review

The structure used in the one principal and one agent model was first studied by [41] and [20], who used a simplified environment with two or three possible outcomes. Similar models where later developed as in [25] and [30]. In contrast with these papers, I allow for a continuum of outcomes, which permits a more complete characterization of optimal contracts. The closest structure to my model is the one studied by [44] who focuses in optimal distortions of the final decision as a mechanism to encourage more information acquisition. This paper, on the other hand, characterizes the optimal contract as the main incentive mechanism.[1]

The information structure used in this paper is similar to the one used in [60] and [56]. However, the first model is used in a procurement environment where the acquired information is induced to be completely revealed, this is not the case in this paper. In the second one the agents acquire information to learn about their value for an object, not the value for a principal as in our model. A similar information structure is also used in [54]. Nevertheless, their model focus in the optimal actions of a single decision maker when information can be acquired over time, and not in the strategic interaction.

The acquisition of information is also related to bandit problems where an agent can learn about the return of a project by undertaking it as in [45] and [23]; however the information structure is more general in our case and contracts do not depend on the realized signal. [11] study moral hazard in teams over time where the return of a project is unknown and effort determines the rate of arrival of the return. Our setup is different in that individuals invest one time on a signal before deciding to undertake the risky project.

The next section introduces a principal agent setup with a single agent acquiring private

[1] [15] also study the motivation of information acquisition, however in their setup the acquired information is observable.

information to give intuition about the optimal contracts on a simplified framework. The third section discusses how to implement the contract in different real-world applications. In the last section I conclude.

1.2 Principal-Agent Problem

Consider the case of a risk neutral principal (she) who has to decide between a safe project with known returns and a risky project with unknown returns. The principal can hire a risk neutral agent (he) to acquire information and recommend him which project to pursue.[2] The information gathered by the agent (if any) is unobservable to the principal. Hence, this is a problem that involves hidden actions since the acquisition of information is not observable to the principal, but also there is a hidden information problem since the realized signal when the agent decides to acquire information is also not observed by the principal. The contract designed by the principal is assumed to be a function of the realized final outcome and the chosen project. It is also assumed that both individuals have limited liability.

1.2.1 Model

There are two available projects that cannot be pursued simultaneously. There is a safe project with known net return $y_s > 0$. There is also a risky one whose return $y_r \in [0, \overline{y}]$ is unknown, with $\overline{y} > y_s$. Let both individuals have the same nondegenerate prior belief $g(y_r)$ over the unknown return with finite mean μ_0.

Before taking the decision, the agent can exert effort and generate information about the risky project by acquiring a signal $x \in \mathbb{R}$ at a cost c. This cost can be associated with

[2]The individual can be in fact risk averse or risk lover, just let the returns perceived by the agent be measured in utils and let the agent maximize a Von Neuman-Morgenstern utility function

the cost of running trials or the disutility of effort. In a context of bandit problems, where signals are the same returns of the risky project, the fixed cost can also be thought as the ex ante expected return of the risky project, μ_0, and reflects the fact that individuals are initially pessimistic about it. The agent can also shirk, in which case no signal is generated and there will be no cost.

Let the conditional pdf and cdf of the signal x be denoted by $f(x|y_r)$ and $F(x|y_r)$, respectively. Assume these are differentiable with respect to x. Similarly, the unconditional pdf and cdf will be denoted by $f(x)$ and $F(x)$, respectively. Let the signals be ordered in the likelihood ratio sense: a signal x is more favorable than signal x' if the posterior distribution $g(y_r|x)$ first order stochastically dominates the posterior distribution $g(y_r|x')$.[3] These type of problems are known as monotone ones and were first studied by [38].

Since signals are ordered, the posterior mean will be a monotone increasing transformation of the signal. Thus the distribution of the posterior mean will be a transformation of the distribution of the signal. Without loss of generalization, let $x = \mathbb{E}_{y_r}[y_r|x]$ be the posterior mean of the risky project. Hence, we can let $\underline{x} = 0$ and $\overline{x} = \overline{y}$.[4]

The acquisition of the signal and its content are privately known by the agent. The only observable variables for the principal are the chosen project and the final return of the project y_s or y_r. Thus the principal designs the optimal wage to be paid to the agent as a function of these variables, that is he chooses the safe wage $w(y_s) = w_s$, and the risky wage schedule $w(y_r)$. It will also be assumed that individuals have a limited liability constraint. The wage for the agent cannot be lower than 0 and the principal cannot pay more than the return he receives. Formally, optimal wages must satisfy $0 \leq w(y) \leq y$. Given that a project

[3]Equivalently, $f(x|y_r)$ is log supermodular

[4]To have an interesting problem we need that $F(y_s) > 0$, otherwise it is optimal to always choose the risky project.

j is chosen and a signal has been acquired, the payoff for the principal is given by $y_j - w(y_j)$ and the payoff for the agent is $w(y_j) - c$.

The game consists of two stages. In the first stage the principal designs a payment schedule and makes a take it or leave it offer to the agent. The agent accepts or rejects the contract. If she accepts the contract, she decides whether to acquire or not a signal, which is privately observed by her. In the second stage, the agent updates her beliefs and chooses which project to pursue. Finally a return y is realized and the principal pays to the agent the contracted wage $w(y)$.

1.2.2 First Best

Suppose the acquired information by the agent is observed by the principal, also assume there are no limited liability constraints. If the principal decides to hire the agent, she will face the following problem:

$$\max_{w(y_r),w(y_s)} \mathbb{E}_x \left[\max_{j(x)\in\{s,r\}} \mathbb{E}_{y_{j(x)}} \left[y_{j(x)} - w\left(y_{j(x)}\right) |x \right] \right]$$

$$\text{s.t } \mathbb{E}_x \left[\mathbb{E}_{y_{j(x)}} \left[w\left(y_{j(x)}\right) |x \right] \right] - c \geq \underline{u}$$

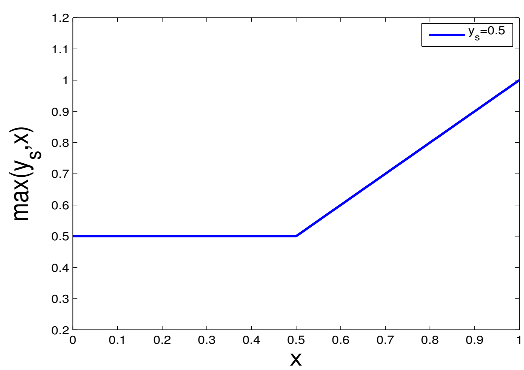
where $j(x)$ is the project chosen when x is realized. The constraint assures the agent will accept the contract by making sure his expected utility is greater than or equal to an outside option $\underline{u}$. The first best can be obtained by either a constant payment from the principal to the agent equal to the cost, or by selling the agent the returns from the project. Either of the alternatives lead us to the following indirect utility when the agent is hired:

$$\mathbb{E}_x\left[\max_{j(x)\in\{s,r\}}\mathbb{E}_{y_{j(x)}}\left[y_{j(x)}|x\right]\right]-c-\underline{u}$$

Since there are two stages, we proceed to solve the individual's problem using backward induction. That is, I will first determine which project is going to be chosen given the information acquired. Then, I will characterize when the principal decides to hire an agent as a function of y_s and c.

The individual will choose the risky project if $x > y_s$, thus the payoff of the second period is given by $\max\{x; y_s\}$. Note this is a convex function of x.

Figure 1.1: **Utility in second period**



The value of experimentation is defined as the ex ante expectation of the utility in the second period, that is

$$U(y_s)=\mathbb{E}_x\left[\max\{x; y_s\}\right]$$

From the previous properties we can prove the following lemma:

Lemma 1 *The value of experimentation* $U(y_s)$ *is strictly increasing and convex in* y_s*, and strictly greater than* $\max\{\mu_0; y_s\}$

The result states that information is always valuable in this setup. However, since information is costly, the principal may not want to hire the agent to acquire information. At the beginning of the first period a principal will choose to hire an agent if and only if $c \leq \hat{c}$, where $\hat{c}$ is defined by

$$U(y_s) - \hat{c} - \underline{u} = \max\{y_s; \mu_0\} \tag{1.1}$$

Let the maximized objective function of the principal be denoted by $V(y_s, c) = \max\{U(y_s) - c - \underline{u}; y_s; \mu$ This function is also strictly increasing and convex in y_s. The next proposition characterizes when the principal decides to hire an agent as a function of y_s.

Lemma 2 *The principal decides to hire the agent when* $y_s \in (a_c, b_c) \subseteq (0, \overline{y})$*, where* $\mu_0 \in (a_c, b_c)$*. Moreover, such interval is decreasing in* c*, that is* $(a_c, b_c) \subset (a_{c'}, b_{c'})$ *for any* $c < c' < \hat{c}$*, with* $(a_0, b_0) = (0, \overline{y})$ *and* $a_{\hat{c}} = \mu_0 = b_{\hat{c}}$*.*

Even if beliefs are relatively pessimistic the individual decides to acquire information because of the potential gain represented by the value of experimentation. The lower is the fixed cost c, the greater is the interval over which the principal decides to hire the agent. Furthermore, if there is no fixed cost, the principal will always decide to hire an agent to collect information.

1.2.3 Constrained Efficiency

Now suppose the principal does not observe the effort of the agent, nor the information gathered by the individual. Also assume individuals have limited liability as described before. In this context a fixed wage will not induce any effort from the agent. Therefore the principal must provide incentives to the agent by imposing more risk in her payoff, and by distorting the choice of the best project. This distortion indicates that the first best will not be attained. The constrained efficient problem for a principal who decides to hire an agent is the following:

$$\max_{w(y_r),w(y_s)} \mathbb{E}_x \left[\max_{j(x)\in\{s,r\}} \mathbb{E}_{y_{j(x)}} \left[y_{j(x)} - w\left(y_{j(x)}\right) |x\right]\right] \tag{P}$$

subject to

$$\mathbb{E}_x \left[\mathbb{E}_{y_{j(x)}} \left[w\left(y_{j(x)}\right)|x\right]\right] - c \geq \underline{u} \tag{IR1}$$

$$\mathbb{E}_x \left[\mathbb{E}_y \left[w\left(y_{j(x)}\right)|x\right]\right] - c \geq \max\left\{\mathbb{E}_y \left[w\left(y_r\right)\right], w_s\right\} \tag{IC1}$$

$$\mathbb{E}_{y_{j(x)}} \left[w\left(y_{j(x)}\right)|x\right] \geq \mathbb{E}_{y_{-j(x)}} \left[w\left(y_{-j(x)}\right)|x\right] \text{ for all } x \tag{IC2}$$

$$0 \leq w\left(y_j\right) \leq y_j \text{ for } j = r, s \tag{LL}$$

Equation (IR1) is the same individual rationality constraint as before. Equation (IC1) is the incentive compatibility constraint that ensures the agent will exert effort. It states that the agent's utility when he exerts effort is greater than the expected utility when he does not, in which case he picks the project that gives her the greatest ex-ante expected wage. Equation (IC2) is another incentive compatibility constraint to make sure the agent chooses the project that is more convenient to the principal. Here $-j(x)$ denotes the project the agent has not chosen. The last equation represents the limited liability constraint.

This problem is hard to solve because the second incentive compatibility constraint involves a continuum of restrictions. However, equation (IC2) can be reduced to only one constraint when optimal wages for the risky project are monotone nondecreasing. Since signals are ordered, a posterior generated by a signal x first order stochastically dominates any posterior generated by any less favorable signal x'. Therefore a less favorable signal implies that the expected wage is lower.

Since distributions are continuous in x, there must exists a cutoff x_e such that the expected wage when the risky project is chosen given such signal is equal to the safe wage. Any more (less) favorable signal than x_e implies the agent will choose the risky (safe) project and that the constraint will not be binding. Formally, there exists cutoff x_e such that:

$$\int_0^{\overline{y}} w\left(y_r\right) g\left(y_r|x_e\right) dy_r = w_s$$

Using this fact, the ex-ante agent's utility can be rewritten as:

$$\int_{x_e}^{\overline{y}} \int_0^{\overline{y}} w\left(y_r\right) f\left(x|y_r\right) g\left(y_r\right) dy_r + w_s F\left(x_e\right) - c$$

By integrating the possible signal values when they are higher than the cutoff we obtain the relevant object to provide incentives: $1 - F\left(x_e|y_r\right)$. This is the probability that the agent chooses the risky project when he exerts effort. Let us for now assume that wages are nondecresing, and in the next proposition I will prove that is the case. The problem can be reexpressed as:

$$\max_{w(y_r),w(y_s)} \int_{\underline{y}}^{\overline{y}} \left(y_r - w\left(y_r\right)\right)\left(1 - F\left(x_e|y_r\right)\right) g\left(y_r\right) dy_r + \left(y_s - w_s\right) F\left(x_e\right) \qquad \text{(P}*\text{)}$$

subject to

$$\int_0^{\overline{y}} w(y_r)(1 - F(x_e|y_r))\, g(y_r)\, dy_r + w_s F(x_e) - c \geq \underline{u} \tag{IR*}$$

$$\int_0^{\overline{y}} w(y_r)(1 - F(x_e|y_r))\, g(y_r)\, dy_r + w_s F(x_e) - c \geq \max\{\mathbb{E}_y[w(y_r)], w_s\} \tag{IC1*}$$

$$\int_0^{\overline{y}} w(y_r) \frac{f(x_e|y_r)\, g(y_r)}{f(x_e)} dy_r = w_s \tag{IC2*}$$

$$0 \leq w(y_j) \leq y_j, \quad \text{for } j = r, s \tag{LL}$$

Let λ, δ_r, δ_s, and ϕ be the Lagrange multipliers for the first three constraints. The distinction between δ_r and δ_s arises because the expected wage from the risky project does not need to be necessarily equal to the safe wage. Since the problem is linear on the wages, the optimal wages are determined by a bang-bang solution that is bounded by the limited liability constraint. After rearranging the derivative with respect to wages we obtain

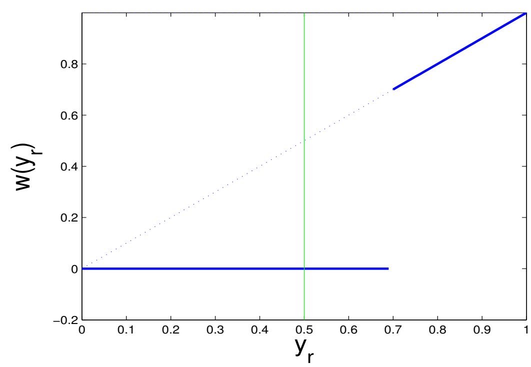
$$-1 + \lambda + \delta_r \left(1 - \frac{1}{1 - F(x_e|y_r)} + \frac{\delta_s}{\delta_r}\right) - \phi \frac{f(x_e|y_r)}{(1 - F(x_e|y_r))\, f(x_e)} \tag{1.2}$$

Whenever this condition is positive, the wage will be set to the upper bound; if it is negative, then the optimal wage is zero. Using the structure of the signals and the latter equation we can indeed prove the that wages are nondecreasing in the following proposition. Figure 1.2 illustrates the contract.

Proposition 3 *If $f(x|y_r)$ satisfies the likelihood ratio order then the optimal wage schedule $w(y_r)$ is monotone and is characterized by a cutoff z such that*

$$w(y_r) = \begin{cases} y_r & \text{if } y_r \geq z, \\ 0 & \text{otherwise.} \end{cases} \tag{1.3}$$

Figure 1.2: **Optimal Contract**



The cutoff z is given by the value of y_r such that condition (1.2) is equal to zero. Wages are thus monotone if for any greater (lower) y_r the derivative is positive (negative). Thus, showing that the derivative is increasing in y_r implies that wages are monotone.

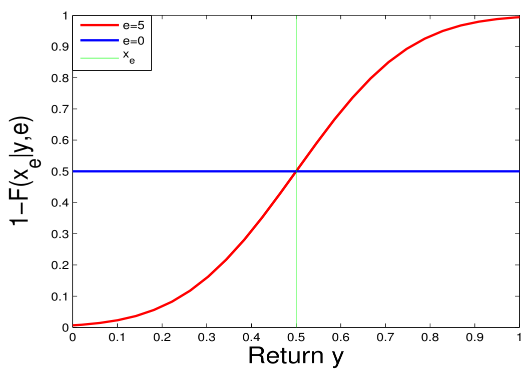
Condition (1.2) has four elements. The first one is the marginal cost to the principal of increasing the wage. The second one is the benefit from relaxing the IR constraint which might be 0 if such constraint is not binding, that is if $\underline{u} < w_s$. This case arises when the agent gets information rents for the private information.

As it is common in moral hazard problems, the third term is related to a likelihood ratio. It is related to the probability of undertaking the risky project when no information has been acquired relative to that same probability when the information has been acquired. Whereas the first one is independent of y_r, the second one is increasing in y_r since the likelihood ratio order induces first order stochastic dominance.

There are three possible cases. If w_s is greater than $\mathbb{E}_y [w (y_r)]$, implying that $\delta_r = 0$, then it would become obvious that the agent exerted effort whenever the risky project is chosen. In this case the problem simplifies to one of just hidden information where incentives must be given to encourage the adoption of the risky project. On the other hand, if $w_s < \mathbb{E}_y [w (y_r)]$, then an agent choosing the safe project would suggest the acquisition of information. Thus, there will always be a punishment for choosing the risky project, but such punishment will decrease the greater is the return. When y_r approaches to $\overline{y}$, there will be no punishment.

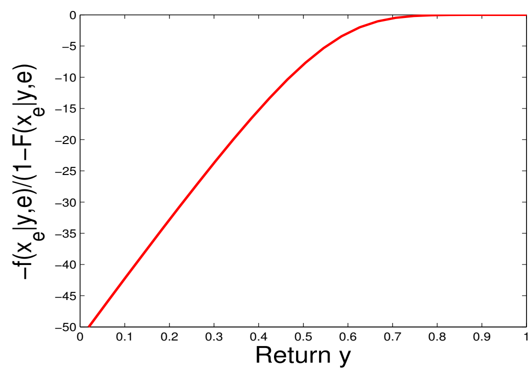
When $w_s = \mathbb{E}_y [w(y_r)]$, the principal is minimizing the cost of inducing acquisition of information. To see this note the similar structure of restriction (IC1) with equation (1.1) and its subsequent properties. In this case the agent will be indifferent between the safe and the risky project when he does not acquire information and it will appear as if he randomizes his decision and chooses the risky project with probability $\frac{\delta_r}{\delta_r+\delta_s}$. Then the principal can reward the agent for risky returns that induced a greater probability of choosing the risky project when the information was acquired, otherwise the agent will be punished. As argued before, such likelihood ratio will be monotone increasing as it is shown in Figure 1.3.

Figure 1.3: **Likelihood Ratio**



The final component is associated to the hidden information problem and is represented by the hazard rate $\frac{f(x_c|y_r,e)}{1-F(x|y_r,e)}$. The hazard rate in this context is interpreted as the probability of having observed the cutoff signal given that the risky project was chosen. Intuitively, the principal wants to make sure the right project is chosen, which implies that, in the limit, she wants to penalize agents who chose the risky project when they just observed the cutoff signal. Given the MLRP condition of the signals with respect to the return, this hazard ratio is monotone decreasing with respect to the return (see [8] and Figure 1.4 for an example). In other words, a lower observed return increases the chances that the agent observed the cutoff signal. Therefore, this effect also suggests that compensation should be increasing in y_r.

Figure 1.4: **Negative Hazard Rate**



In sum, a monotone contract encourages information acquisition and the partial revelation of information. On one hand, paying more for higher outcomes induces the acquisition of information since this increases the probability of choosing the best project. On the other hand, paying more for higher outcomes induces the agent to choose the best project once the information has been revealed since a higher signal is associated to a higher return. Thus the optimal wage schedule creates the incentive for the agent to adopt the risky technology after observing favorable signals.

The safe wage is set according to constraint (IC2$*$). Formally, choosing w_s is equivalent to choosing x_e. The first order condition with respect to x_e is simplified to:

$$y_s - x_e - \phi \frac{\partial \mathbb{E}\left[w\left(y_r\right)|x_e\right]}{\partial x_e} = 0$$

Raising the threshold will yield a marginal benefit of y_s since the safe project will be pursued more often. However, it also generates a cost of $x_e = \mathbb{E}\left[y_r|x_e\right]$, the return of the risky project at the threshold, and a cost of increasing the expected wage when such threshold increases. The last cost arises because the safe wage must increase to generate the appropriate incentives to the agent to choose the desired project. [5]

This condition implies that the optimal decision is always distorted in the constrained efficient solution in favor of the risky project, $x_e < y_s$. This distortion is purely generated by the hidden information problem. Note that constraints (IR$*$) and (IC1$*$) remained constant after a change on x_e precisely because of constraint (IC2$*$). Since it is costly for the principal to induce the agent to reveal the information through the risky wage, she favors the decision that reveals the uncertain state.

This distortion also suggests that the first best is never attained. To avoid the distortion the risky wage should be constant. Although the hidden information would be solved in this case, the moral hazard will persist since information will have no value to the agent. The first best is also never attained because the principal will hire the agent only for c strictly lower than $\hat{c}$, where $\hat{c}$ was defined in equation (1.1). To see this just note that the value of experimentation decreases for the principal since she will only appropriate risky returns lower than z, thus her payoff is no longer convex.

Finally note that z must be greater than y_s. If this is not the case then the principal will never want to hire the agent since she is better off by pursuing the safe project. This result suggests that the agent should be rewarded for pursuing the risky project only when its return is significantly better than the safe return, not just marginally better.

[5]Remember that wages are monotone increasing, implying that a higher threshold generates a higher expected wage.

1.2.4 Debt Contracts

The optimal contract found in the previous subsection is not continuous. In particular, the payment for the principal is not monotone since any return greater than the threshold will yield him zero profit. As argued by [37], this type of contracts could be manipulated by either the principal or the agent if any of them can affect the return before the contract is paid. For example, the principal would have incentive to sabotage the risky project by burning profits in excess of the threshold. Similarly, the agent would have an incentive to inflate profits by borrowing money and "revealing" a higher apparent profit to the principal.

In order to prevent this behavior, a monotonicity constraint must be imposed, thus modifying the limited liability constraint:

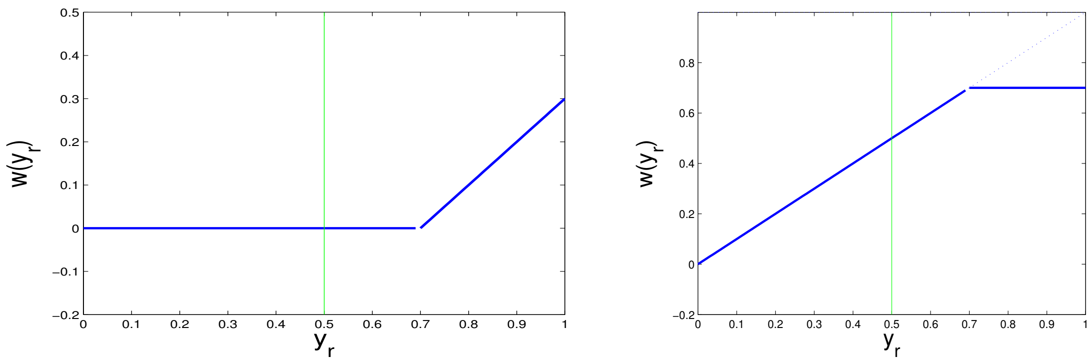
$$w\left(y_r-\epsilon\right)\leq w\left(y_r\right)\leq w\left(y-\epsilon\right)+\epsilon$$

The same argument in the preceding section applies, and the optimal monotone contract will be option like, with an strike price z_0 greater than the safe return:

$$w\left(y_r\right)=\max\left\{0,y_r-z_0\right\} \tag{1.4}$$

This type of contract was first obtained by [37] using a principal-agent setup. In his paper the argument relies on the standard assumption that a greater effort generates better distributions of profits in the likelihood ratio sense. Therefore an optimal contract rewards the agent for higher profits and punishes him otherwise.

Figure 1.5: **Optimal Monotone Contract**



(a) Payment to Agent (b) Payment to Principal

The difference with our setup is that effort does not lead to such ordering on the distribution of returns. In fact, effort does not directly influence the observed returns. The returns for either project are given ex-ante, the problem is that individuals are uncertain about the risky return. Effort in our context lets the agent take a more informed decision and that is why the agent is rewarded for good outcomes.

1.3 Implementation

The studied setup closely resembles the interaction between a CEO and the shareholders of a firm. The shareholders hire a CEO to take decisions concerning the future of the company. It is usually the case that the CEO has more expertise than the shareholders in making such decisions, or at least is more efficient at gathering information related to such decisions. Decisions can range from acquisitions to the marketing of new products, whose main characteristic is the uncertainty of their return. Such uncertain returns will be reflected in the value of the firm and thus affect the shareholders payoff. The effort exerted by the CEO to acquire information is not observable. It is also common that if information has been acquired by the CEO, it is not (completely) observed by the shareholders.

The optimal contract derived here suggests that stock options and profit sharing compensation are optimal ways to motivate the acquisition of information and its revelation. Stock option programs give workers the right to buy company's shares at a fixed price for a given period of time. These will only be exercised if the market price is higher than the strike price originally agreed to. Stock options are thus used as a long- term motivator and the employee is constrained on exercising the option only after their performance has been (partially) observed. Likewise, profit sharing is also used as a long-term motivator where individuals are entitled to a percentage of the profits of a firm after a given period. To implement the contract the firm could set a threshold on the profits such that the CEO can only claim his share if profits are greater than such level.

The constrained efficient solution also explains why it is optimal to pursue risky projects even when the information suggests the safe project is better. This feature is in close connection to the debate on excessive risk taking of CEOs. Although this is inefficient ex-ante, it is a valuable strategy because it helps to reveal the uncertain state of nature. This is also related to the result obtained by [45] where early failure is not punished and long-term success is rewarded.

The environment also describes situations that involve technology adoption. The return of new technologies is usually uncertain until they are tried. For example, it has been documented the lack of adoption of high yielding variety seeds in developing countries because of the uncertainty that farmers must bear ([26]). A social planner interested on the adoption of new technologies but unable to observe the effort (or number of trials) performed by the farmers and their subsequent results, could use an option-type contract to encourage farmers.

The optimal contract might be also interpreted as a patent policy to encourage innovation. Innovation is usually thought as a process whose return is proportional to the amount of R&D expenditures. However, one could think of such expenditures as being a measure of the precision of the signals that emerge from the trials. If a great quantity of money is invested and the trials suggest that the new product would not work, a good decision is to stop pursuing such project. On the other hand, if not enough money has been invested, a bad trial does not necessarily imply is not a good project.

Following such interpretation, a social planner who wants to increase the levels of innovation in a society but cannot observe R&D expenditures and their outcomes, should use the optimal contract derived here to provide incentives for innovation. The contract can be interpreted in terms of a patent. It suggests that patents should only be given if it is shown that the new technology is significantly better than the previous one, and not for marginal improvements. This result cannot be interpreted as a restriction on the use of the new technology as often happens with patents. In other words, the optimal contract does not allocate the property rights of the new technology. On the contrary, it encourages the adoption of the new technology by all the population, while rewarding innovators with the surplus they generated, suggesting an optimal pricing policy.

1.4 Conclusions

This paper studies the problem faced by a risk-neutral principal who must hire a risk-neutral agent to take a decision between a safe project with known return and a risky one with unknown return. The agent can acquire costly information about the new project before taking the decision of whether to pursue it or not. The acquisition of information and its content are unobservable to the principal, thus the problem features moral hazard and hidden information frictions.

The main result is that the wage schedule when the risky project is pursued must be monotone increasing. The agent will be rewarded with the whole output only if the return is significantly better than the safe return, otherwise he will be paid nothing. Thus the risk of the decision is imposed to the agent. It has been also shown that decision making is distorted in favor of the risky project. That is, it is optimal to pursue the risky project when a signal reveals that is slightly worse than the safe one.

The optimal contract resembles restricted stock in an environment where shareholders hire a CEO to make decisions. In this case the CEO will be given firm shares when he decides to undertake the risky project and its returns are significantly greater than the returns from the previous project. When the optimal contract is constrained to be continuous it resembles an option with a strike price greater than the return of the safe project. The model also provides explanation as to why excessive risk taking is optimal, where excessive risk taking is understood as pursuing the risky project even when it seems to be worse than the safe one.

The conclusions of the model are robust when there is a finite number of possible projects. However, the same framework does not apply when the decision belongs to a continuum as in [44]. In this case the optimal contract is analogous to one where the gathered information is completely revealed. This is a natural extension of the proposed model that we leave for further study. Another useful extension is the characterization of the optimal contract when the agent is risk-averse as well as in a dynamic setup where reputation for undertaking good projects can be built.

1.5 Appendix

Proof for Lemma 1. To show that $U(y_s) > \max\{\mu_0, y_s\}$ integrate by parts the value of experimentation to obtain

$$
\begin{aligned}
U(y_s) &= y_s + \int_{y_s}^{\overline{y}} (1 - F(x))\, dx \\
&= y_s + \int_{0}^{\overline{y}} (1 - F(x))\, dx - \int_{0}^{y_s} (1 - F(x))\, dx \\
&= \mu_0 + \int_{0}^{y_s} F(x)\, dx \qquad (1.5)
\end{aligned}
$$

Also note that

$$\frac{\partial U(y_s)}{\partial y_s} = F(y_s) > 0$$

where the strict inequality comes from footnote 4. Convexity is also easily obtained since the second derivative is the probability of having a signal equal to y_s.

■ **Proof for Lemma 2.**

First note that $U(0) = \mu_0$ and $U(\overline{y}) = y_s$. Therefore $U(y_s) - c$ will cross at most once each of the outside options. It could cross once the constant μ_0 from below since it is increasing in y_s. It could cross once y_s from above since its first derivative with respect to y_s is between 0 and 1. This in turn implies that $U(y_s)$ is farther from $\max\{\mu_0; y_s\}$ precisely when $\mu_0 = y_s$.

Since $U(y_s) - c$ is linear in c, there exists a $\hat{c}$ such that $U(y_s) - c = \mu_0 = y_s$. Thus, for any $c < \hat{c}$, there exists $a_c, b_c \in (0, \overline{y})$ such that $U(y_s) - c > \mu_0$ for any $y_s > a_c$, and $U(e^*) - C(e) > y_s$ for any $y_s < b_c$. Obviously it must be the case that $\mu_0 \in (a_c, b_c)$. Note that a_c and b_c are increasing and decreasing in c, respectively, precisely because the function crosses from below and above each of the corresponding outside options. Finally, for any $c > \hat{c}$, the interval is empty and the principal never experiments.

■

CHAPTER 2

Endogenous Employment and Incomplete Markets

2.1 Introduction

The unemployment rate for college graduates is lower than for non-college graduates (4.4% versus 9.6% in 2011 according to the Bureau of Labor Statistics), as well as the median duration of spells of unemployment (2.6 months for less than high school graduate, 2.4 for a high school graduate and 1.9 months for individuals with at least some college). It is also true that college education is usually obtained by richer households, which creates a stronger tension towards a more unequal distribution of wealth.

However, recent empirical papers have provided evidence on the negative effect that wealth has on the probability of employment once a set of control variables, including human capital, are used [6,9,10,58]. On the theoretical side, such association was rationalized by [42] in a search model with savings where effort must be exerted over time to increase the chance of being employed.

The above observations suggest a rather complex relationship between assets and the probability of employment, which seems to be negative in the short-run put positive in the long run. The purpose of this paper is to build a model to explore the joint role of effort and human capital investment as non-market mechanisms used by individuals to deal with their idiosyncratic risk. The analysis provides potential welfare implications for combining public policies related to unemployment insurance and subsidies to education to improve the wealth distribution and the long-run unemployment.

I develop a dynamic stochastic general equilibrium with heterogeneous agents that builds on the framework proposed by [36] and [5]. Effort and human capital are variables determining the transition dynamics between states, whereas asset holdings are used directly to smooth consumption. Effort is modeled as a flow variable that has to be chosen every period to maintain a positive probability of being employed, thus following the literature on unemployment insurance (see for example [34] and [61]). This can be seen as search effort when the individual is unemployed, or effort in the job when the agent is employed. We assume the level of effort required in the latter case is more effective that the one when the agent is unemployed. This assumption matches with empirical data that has been studied in search models and emphasize the role of the depreciation of human capital during unemployment ([3]; [55]).

On the other hand, human capital is a binary stock variable that can be acquired when the individual is born. It improves the efficiency of effort when looking for a job or maintaining it. Although human capital has been usually studied as a mechanism to increase earnings, previous empirical work has also pointed out the effect of human capital on employment transitions. For example, [14] estimate the effect of training on the probability of employment for the 1976 cohort of adult male participants in the Comprehensive Employment and Training Act (CETA). They found that the effect is positive, even for people who is already employed. [29] also found that participation of women in private training programs increases both the frequency and duration of employment spells.

Although there exists some selection on the individuals that attend to college due to differences in abilities for example, I abstract from this issue. I assume agents are homogeneous in this dimension and that human capital does not affect the income when employed. Even with this simplication the model does a good job on replicating the wealth distribution.

As it is usual in this literature, asset holdings are restricted to be greater than a lower bound to prevent situations where individuals get indebted forever. This lower bound is used to model a financial friction usually found in reality, and is calibrated accordingly. An upper bound arises naturally from the optimal decisions and the fact that the interest rate is lower than the rate implied by the intertemporal discount factor. This discourages individuals from accumulate forever their asset holdings.

The role of the asset holdings in our model is similar to the one played in previous literature. When the individual is employed she accumulates assets, while she decreases her holdings when unemployed. Therefore, it keeps track of the employment history the individuals have had. However, assets also have a bequest motive in this model. Individuals die with an exogenous probability and newborns inherit the previous wealth. Given the assumptions, if the cost of attend college is sufficiently large, only sufficiently rich born individuals invest in human capital. This generates pressure towards more inequality.

On the other hand, effort has an inverse relationship with assets. If an individual becomes unemployed and has sufficient savings, she will not exert too much effort to find a job and instead use the savings to smooth consumption. However, the ability of the assets to smooth consumption loses importance when they are close to the debt limit. At that point effort plays a major role by increasing the likelihood of being employed next period.

In the stationary distribution most of the individuals will hold a small negative credit balance, while few of them will have positive savings. This means that most of the individuals combine both channels to smooth consumption rather than relying in one of them. As a consequence, the resulting stationary distribution of wealth is much closer to the real one than the wealth distributions obtained by previous studies. Papers that focus on market mechanisms to alleviate risk usually generate wealth distributions negatively skewed since precautionary savings are the only channel to smooth consumption. On the other hand, empirical papers have shown that only the top deciles have positive savings, while most households hold some degree of debt [63].

Idiosyncratic shocks and consumption smoothing has been largely studied in the literature. Models of incomplete markets and heterogenous agents have been used to explain the risk premium [36], the benefits of insurance [31], optimal fiscal policy [32], and the distribution of income [5, 33, 39], among others. The common characteristic of these models is that they use mechanisms affecting the budget constraint to smooth consumption. These mechanisms are usually identified with assets holdings (or credit balances), capital, or savings. However, the labor transitions are always specified exogenously.

Besides the theoretical contribution made by [42] on endogenous transitions, other calibrated models of search with savings include [2], [57] and [28]. However, the inclusion of human capital and the characterization of the wealth distribution are new in our model.

The organization of the paper is as follows. The next section describes the model and the third section defines the equilibrium in this scenario. I then describe the performed numerical exercise, while section 5 devotes attention to its computation. In section 6 we show the results and its implications. The last section concludes.

2.2 Model

Consider an exchange economy with a continuum of agents with total mass equal to one who face idiosyncratic risk. There are two commodities: one perishable consumption good c and asset holdings a. Each agent receives an stochastic endowment s_t at the beginning of each period. Assume the endowment can take two possible values $s_L < s_H$, which are usually associated with unemployed/employed status, respectively.

Effort $e > 0$ is made in order to increase the probability of having a good endowment (state) next period. The probability of being employed next period also depends on whether the agent has a college degree or not, h_H or h_L, respectively. The probability in period t is defined as $\Pr(s_{t+1} = s_H|s_t, h) = P(e_t; s_t, h)$, which is an increasing concave function of the effort with $P(0; s, h) = 0$ and $\lim_{e\to\infty} P(e; s, h) = 1$. According to the empirical literature, assume that effort to remain employed is more effective than the effort to become employed when previously unemployed, and effort is also more effective when the individual has a college degree. Formally, $P(e_t; s_H, h) > P(e_t; s_L, h)$ for all h, $P(e_t; s, h_H) > P(e_t; s, h_L)$ for all s. Finally let the probability be supermodular in e, s, and h.

Individuals discount future at rate β and survive next period with probability δ. When an individual dies it is replaced by an unemployed newborn. The newly born agent inherits previous wealth and decides whether to obtain a college degree or not at a cost ϕ. Agents are altruistic and maximize lifetime utility of the household. Each individual derives instantaneous utility from consumption and effort according to an additive separable utility function $u(c) - e$ that is strictly concave and satisfies Inada conditions. Separability can be obtained assuming the existence of lotteries and simplifies the analysis importantly [42]. The fact that the disutility of effort is linear is just an innocuous normalization.

Each agent is able to smooth her consumption by holding a single riskless asset. This asset entitles the individual to receive one unit of future consumption for each unit of asset whose price is $q > 0$. The amount of claims held must remain above the limit $a_{\min}$, a restriction that represent the financial friction faced by individual in addition to the incompleteness of the markets. Therefore, the budget constraint faced by an individual who holds a claims, has a current endowment s, and chooses consumption c and future claims a', is given by

$$c + qa' \leq s + a \tag{2.1}$$

The agent's problem can be represented in recursive formulation as

$$\begin{aligned} v\left(a, s, h; q\right) = \max_{c,e,a'} \{ & u\left(c\right) - e + \beta\delta\left[P\left(e; s, h\right) v\left(a', s_H, h; q\right)\right. \\ & \left. + \left(1 - P\left(e; s, h\right)\right) v\left(a', s_L, h; q\right)\right] + \beta\left(1 - \delta\right) v_0\left(a'; q\right) \} \end{aligned} \tag{2.2}$$

subject to (2.1), $c \geq 0$, $e \geq 0$, and $a' \geq a_{\min}$; and where

$$v_0\left(a'; q\right) = \max\left\{v\left(a', s_L, h_L; q\right); v\left(a' - \phi, s_L, h_H; q\right)\right\}$$

This problem is well defined since $v\left(a, s, h; q\right)$ will inherit the concavity properties of $u\left(\cdot\right)$, while also satisfying discounting and monotonicity (see [59]). On the other hand, the functional v_0 can be replaced without loss of generalization by its least concave function. Therefore, the first order conditions are necessary and sufficient, and the optimal decision rules $c(a,s,h;q)$, $e(a,s,h;q)$, and $a'(a,s,h;q)$ are given by

$$\begin{aligned}
1 &\geq \beta P_e(e;s,h)\left[v(a',s_H,h;q) - v(a',s_L,h;q)\right], \\
\text{with equality if } e > 0 & \\
u_c(c)\,q &\geq \beta\delta\mathbb{E}\left[u_c(c')\,|e,s,h\right] + \beta(1-\delta)\frac{\partial v_0(a';q)}{\partial a'}, \\
\text{with equality if } a' &> a_{\min} \\
c + qa' &\leq s + a
\end{aligned}$$

The first condition shows the tradeoff between the marginal disutility and the expected marginal benefits of exerting an effort. This condition is similar to the one obtained in the optimal unemployment insurance literature. Using the separability of the utility function we can prove the following lemma:

Lemma 4 *Effort is a decreasing function of assets*

The intuition behind this result relies on the fact that the difference on the value function for employed and unemployed people is decreasing in assets, formally $v(a',s_H,h;q) - v(a',s_L,h;q)$ is decreasing in a by supermodularity. In other words, it is less important for rich households whether they are employed or unemployed since they can use their assets to smooth consumption. Therefore the role of effort becomes less important. On the other hand, poor households cannot incur in more debt when they are close to the debt limit. Hence their current state generates great differences in their maximized utility so effort becomes crucial in increasing the probability of being employed.

The difference $v(a',s,h_H;q) - v(a',s,h_L;q)$ is also decreasing in assets by a similar reason. Using this fact and the concavity of the value function on a, we can obtain the following lemma:

Lemma 5 *Let* $A(\phi) = \{a > a_{\min} + \phi : v(a', s_L, h_L; q) < v(a' - \phi, s_L, h_H; q)\}$ *be the set of assets holdings such that a newborn with inherited wealth* $a \in A(\phi)$ *will prefer to acquire a college degree. Then there exists* ϕ_1 *and* ϕ_2 *such that:*

- *For any* $\phi < \phi_1 < \phi_2$*, there exists* $a_1(\phi)$ *such that* $A(\phi) = (a_{\min} + \phi, a_1(\phi)]$*, where* $a_1(\phi)$ *is increasing in* ϕ*;*

- *For any* $\phi > \phi_2 > \phi_1$*, , there exists* $a_2(\phi)$ *such that* $A(\phi) = [a_2(\phi), \infty)$*, where* $a_2(\phi)$ *is increasing in* ϕ

- *And if* $\phi_2 < \phi < \phi_1$*, then* $A(\phi) = [a_2(\phi), a_1(\phi)]$*.*

The lemma states that if the cost of education is sufficiently low, only poor individuals that can afford it will attend college. The reason behind the result is that, when the cost is low, individuals that decide not to acquire education have a greater marginal value for the assets than households that decide to invest. Then if some individual was indifferent between going to college or not, a richer one will definitely prefer to avoid the investment. However, if the cost of college is large the inequality reverses because of the concavity of the value function. Thus the marginal value for individuals that decide to attend college becomes greater and only rich enough individuals will attend college.

The second first order condition is very familiar to the literature that uses asset markets. The limiting behavior of consumption can be characterized by applying the theory of martingales. Let $Z_t = \left(\frac{\beta\delta}{q}\right)^t u_c(c_t) \geq 0$. Therefore, $E_t[Z_{t+1} - Z_t | I_t] = \left(\frac{\beta\delta}{q}\right)^t E_t\left[\frac{\beta\delta}{q} u_c(c_{t+1}) - u_c(c_t) | I_t\right] <$

0, where I_t is the information set at time t, including e_t. The previous expectation implies that Z_t is a supermartingale. Since Z_t is nonnegative, we can apply the supermartingale convergence theorem. This theorem states that Z_t must converge almost surely to a nonnegative random variable [62]; which leads to the following lemma:

Lemma 6 *In equilibrium $\beta\delta < q$*

If $\beta\delta > q$ then Z_t must converge to zero to avoid its divergence. But then this implies that c_t must diverge to infinity. This is obtained by letting the asset holdings go to infinity since the incentives to save are greater than the ones to get more debt. This explosive solution can not be an equilibrium. A similar behavior is obtained if $\beta\delta = q$, see [16] for an exposition.

On the other hand, if $\beta\delta < q$, then Z_t will converge to a nondegenerate nonnegative random variable. This implies that consumption and asset holdings will remain finite, a necessary condition to achieve an equilibrium. In fact, there will be an endogenous upper bound such that no agent would like to save more than such bound (see [43]). The first order condition also implies that agents will save when facing a good shock and spend savings when facing an adverse shock.

It is important to note that optimal decision rules will depend on their state vector (a, s, h) and on the price of claims q. This price will be determined in equilibrium according to a market clearing condition for the asset holdings. The existence of such equilibrium is easy to obtain given the standard properties of the model; however, the equilibrium will not be unique. Since we are interested in the long run interaction in this economy, we focus only on the stationary equilibrium that we describe in the next section.

2.3 Stationary Equilibrium

The equilibrium in an exchange economy is usually defined as policy rules and prices that clear the markets given some aggregate states. However, the market clearing condition is always changing in this dynamic economy given that the distribution of individuals is always moving. Therefore, a definition of a stationary equilibrium is more appropriate in this context. In this definition we focus on market clearing when the distribution of wealth λ is invariant and plays the role of the aggregate variable that depends on the price q.

The law of motion of this state vector distribution is described by

$$\begin{aligned}\lambda_{t+1}\left(a', s', h_H; q\right) &= \Pr\left(a_{t+1} = a', s_{t+1} = s', h_{t+1} = h_H\right) \\ &= \delta \int_{\{a:a'(a,s,h_H;q)\}} \sum_{i=L,H} \lambda_t\left(a, s_i, h_H; q\right) \cdot P\left(e_t; s_t, h_H\right) da_t \\ &\quad + (1-\delta) \int_{a\in A} \sum_{i=L,H} \sum_{j=L,H} \lambda_t\left(a, s_i, h_j; q\right) da\end{aligned}$$

and

$$\begin{aligned}\lambda_{t+1}\left(a', s', h_L; q\right) &= \Pr\left(a_{t+1} = a', s_{t+1} = s', h_{t+1} = h_L\right) \\ &= \delta \int_{\{a:a'(a,s,h_L;q)\}} \sum_{i=L,H} \lambda_t\left(a, s_i, h_L; q\right) \cdot P\left(e_t; s_t, h_L\right) da_t \\ &\quad + (1-\delta) \int_{a\in A^c} \sum_{i=L,H} \sum_{j=L,H} \lambda_t\left(a, s_i, h_j; q\right) da\end{aligned}$$

A stationary distribution is thus defined as a distribution $\lambda(a, s, h; q)$ such that $T\lambda(a, s, h; q) = \lambda(a, s, h; q)$. The existence and uniqueness of the invariant distribution is established using the approach suggested by [35]. Therefore, starting from any initial distribution, a sufficient number of iterations will converge to the invariant one. Moreover, since $a'(a, s, h; q)$ is bounded, the sequence of averaged assets will also converge.

Definition 7 *A stationary equilibrium is defined by policy rules $c(a, s, h; q)$, $e(a, s, h; q)$, and $a'(a, s, h; q)$; a value function $v(a, s, h; q)$; a price q; and a stationary distribution $\lambda(a, s, h; q)$, such that*

- *The policy and value functions solve the agent's problem (2.2)*
- *Markets clear:* $\int_a \sum_{i=1,2} \sum_{j=1,2} a'(a, s_i, h_j; q)\, \lambda(a, s_i, h_j; q)\, da = 0$
- *The stationary distribution $\lambda(a, s, h; q)$ is induced by the policy functions and the endogenous Markov chains generated by $P(e(a, s, h; q); s, h)$.*

The first condition states the optimality of the decisions. The second one defines market clearing for assets, which means that the average holdings in the population must be zero. By Walras Law, if the market of loans is cleared, then the market of the consumption good is also cleared by making average consumption equal to the average endowment. The third condition requires that the distribution of assets remains the same over time. For that we need them to remain finite, this is assured by the lower bound and the fact that $\beta\delta < q$. It also plays an important role that $P(e; s_H, h) > P(e; s_L, h)$.

2.4 Numerical Exercise

We calibrate the model according to the previous literature on heterogenous agents, mainly [36], and unemployment insurance [34]. We first assume the utility function takes the form

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}$$

This is the standard utility function used in this type of problems. According to [52], estimates of the risk aversion coefficient σ are around 1.5. The rest of the parameters are calculated according to periods of 8.5 weeks approximately, that is 6 periods per year. [36] chose this length to match the average duration of unemployment spells of 17 weeks (Bureau of Labor Statistics), which is a underestimation of the current average duration of 21.6, but it fits the 5 year trend. For this the endowments were calibrated to $s_H = 1$ and $s_L = 0.1$, where the last number assumes that individual has access to social programs when he is unemployed. Finally $\beta = 0.99322$ to match an annual discount rate of 0.96, and$\delta = 0.995$ to match the average death rate.

Hugget also specified an exogenous Markov process where $\Pr(s_{t+1} = s_H | s_t = s_H) = 0.925$ and $\Pr(s_{t+1} = s_H | s_t = s_L) = 0.5$. This calibration replicates a coefficient of variation for the annual earnings of 20%, which is close enough to the actual data. It also generates an annual average endowment of 5.3; therefore, we set $a_{\min} = -5$ to simulate the financial friction. This bound generates in equilibrium an annual interest rate between 2.3% and 3.4% in Huggett's calculations and is close to the natural borrowing limit of $-\frac{s_L}{r}$ described by [5].

In order to obtain similar quantitative results, we calibrate our endogenous Markov chain to find similar probabilities. We model the probability of having a high state tomorrow as a cdf of an exponential distribution with parameter $\mu(s_i, h_j) = s_i h_j$, that is $P(e; s_i, h_j) =$

$1 - \exp^{-s_i h_j e}$, where $h_L = 13$ and $h_H = 16$. This parameterization satisfies our initial assumptions of first order stochastic dominance and the ones described by [34] to characterize the optimal unemployment insurance. Moreover, as shown in the next section, the optimal probabilities in equilibrium will wander around Huggett's calibration. Finally, the cost of education is set to $\phi = 4$ to match the average cost of public college relative to average income (see the 2011 report from the College Board). Note this number allows individuals with some degree of debt to invest in education.

2.5 Computation

To find the optimal policy rules we first set a candidate for q, say q_1, belonging to a plausible interval of equilibrium prices. We then use value function iteration to obtain the optimal policy rules. Since all the desired properties of the value function are satisfied, convergence is achieved independently of the initial guess for the value function. To compute the solution we discretize the choice of a, obtain e from its first order condition and consumption from the budget constraint. The grid must be fine enough to achieve smooth policy functions.

As pointed out before, there exists a natural upper bound for a. Optimal future assets for an employed agent start above the 45^0 line (when current assets are negative), and then crosses this line for some positive level of current holdings, say $a_{\max}$. On the other hand, an unemployed agent will always reduce her holdings to maintain her consumption. See Fig. 2 in the appendix for an example of an optimal policy rule for asset holdings.

This shape of the optimal policy implies that $a_{\max}$ plays the role of a fixed point when an agent is always employed. Moreover, it also plays the role of an upper bound since once the agent receives a bad shock she will decrease her assets. Hence, an agent with any initial wealth will converge to the interval $[a_{\min}, a_{\max}]$, and remain there forever. This upper bound can only be computed by experimentation and thus the upper bound of the grid is set large enough to include the fixed point.

After obtaining the optimal decision rules we compute the stationary distribution. To obtain it we simulate an economy of 100000 agents and iterate for 200 periods.[1] The initial distribution of states and assets will not matter for the convergence. We first fix a set of two i.i.d shocks $\epsilon_{i,t}$ and $z_{i,t}$ with a uniform distribution between 0 and 1 for each individual and each period. We then interpolate the optimal decision using the optimal policy rules and the current asset holdings and state. If $z_{i,t} > \delta$ the individual dies and a newborn must decide whether to acquire the college degree or not according to the inherited wealth. If this is not the case then I proceed to compare the first i.i.d shock with the probability associated with the optimal effort and the current state. If $P\left(e\left(a_{i,t}, s_{i,t}, h; q\right); s_{i,t}, h\right) \leq \epsilon_{i,t}$ then $s_{i,t+1} = s_L$, otherwise the agent will be employed.

After the stationary distribution is computed we calculate the excess demand for assets given the initial price q_1. Then we follow Huggett's process of bisection: if the excess demand is positive we increase the price q, if it is negative we decrease it. This algorithm follows the conjecture that the excess demand of assets is negatively correlated with its price. Although this is hard to prove in general, this is the case in the interval we examined, and it has been also true in related papers that follow the same methodology (see for example [36] and [5]). The process continues until excess demand is approximately 0 and the difference of the updated price is less than 0.001.

[1] We also chose a longer horizon without obtaining significant differences.

2.6 Results

Fig. 1 shows the concavity of the value function that permits the contraction to find the fixed point. It also shows how utilities diverge when asset holdings are close to the lower limit, a result that is intuitive after examining the policy rules. The optimal asset policy is shown in Fig. 2 and it follows the behavior described in the previous section. It shows how individuals with low states will decrease their holdings until the lower limit, while individuals with good shocks accumulate holdings until they reach the upper bound. This is a characteristic of the models in this branch of the literature.

In our model we also explore a different non-monetary mechanism used by individuals to alleviate risk. Individuals use effort to increase their probability of being employed next period, especially when their level of assets is approaching its lower limit. The optimal probability of employment conditional on human capital is decreasing on the asset holdings and is lower for unemployed individuals since by assumption is harder to change their status. These probabilities are shown in Fig. 3 and wander around the probabilities calibrated by [36], providing a good approximation of the steady state. They also show how the individual increases them when asset holdings are close to the lower bound.

As a consequence of this optimal strategy for risk bearing, consumption has very little variation across different types of individuals, except for unemployed agents whose asset holdings are close to the lower limit. Fig. 4 depicts the optimal consumption. Fig. 5 shows the excess demand of holdings, which depends negatively on the price. The price of assets that clears the market is 0.9933, which is equivalent to an annual interest rate of 4.1%. This rate is higher than the one obtained by [36] since individuals have more incentive to acquire debt instead of save.

The obtained percentage of individuals with a college degree is 30.6% which is close enough to the real one (29%). The simulated rate of unemployment is 5.5%, consistent with unemployment rates for developed economies.[2] The model generates an unemployment rate of 4.6% and 6% for college and non-graduates, respectively. The generated gap between these two rates is not enough, since these numbers are currently 4.9% and 9.5% for the US. The model is also robust to small perturbations in the parameters. In an alternative scenario where $h_L = 16$ and $h_H = 20$, the rate of unemployment decreases to 5%, which is decomposed on 4.5% and 5.7% for college graduates and non-graduates, respectively.

The distribution of wealth in the stationary distribution differs from the one found by [36] and the one potentially generated by the class of models where consumption can only be smoothed through market mechanisms. These models generate distributions skewed to the left since they must accumulate precautionary savings to deal with their idiosyncratic shocks. In contrast, when the transitions are endogenous, individuals will diversify between the market and the non-market mechanisms.

Fig. 6 shows how the wealth distribution in our model is skewed to the right, approximating better the real wealth distribution [63]. This suggests that in the long run individuals are not afraid of becoming indebted since they have an extra mechanism to smooth consumption. At the end, the incomplete markets partial insurance is successfully complemented by the effort. This result is a consequence of the convexity properties of the sets. In our model it can be traced to the concavity of the probability transition to the employed state, as well as the concavity of the utility function.

Fig. 7 shows the disaggregation of the wealth distribution for college graduates and non-graduates. It is found that distribution for the former is more dispersed and less skewed.

[2]Current unemployment for US is 7.6%, which is higher than the trend observed in previous years.

Thus having college education seems to increase the expected wealth but creates more uncertainty at the same time.

2.7 Concluding remarks

I have studied a model of heterogenous agents who face idiosyncratic risk and smooth their consumption using a riskless assets and non-market mechanisms. Effort is a flow variable that must be exerted every period to obtain or maintan a job, whereas human capital increases the efficiency of effort in obtaining a job and persists until the individual dies. We found that effort and assets have an inverse relationship and it is shown how the investment in a college degree depends on its cost. If the cost is sufficiently high, as it appears to be according to the calibration, then only rich-born agents acquire education.

In the stationary equilibrium agents diversify among these mechanisms and as a result I obtain a distribution of wealth that is not as skewed as the one generated by previous models. In particular, the median individual holds a small negative credit balance and exert a medium amount of effort. This result contrasts with the ones previously obtained where the median individual holds a positive credit balance. Therefore, our framework replicates much better the real distribution of wealth.

The analysis also shows how the distribution of wealth for college graduates is more dispersed and less skewed than the distribution for non-graduates. The model could be used as a benchmark to evaluate the combination of unemployment insurance policies with subsidies for education to improve the wealth distribution. It first suggests how asset holdings could be used as a proxy to unobservable effort, and how education can be used as long-run insurance.

2.8 Figures

Figure 2.1: **Value function**

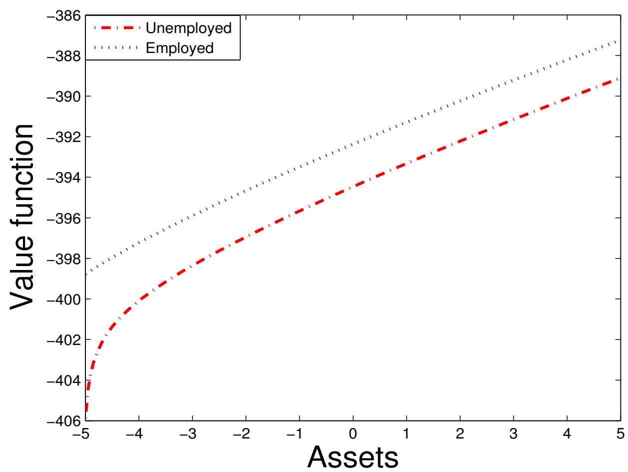


Figure 2.2: **Optimal policy rule for assets**

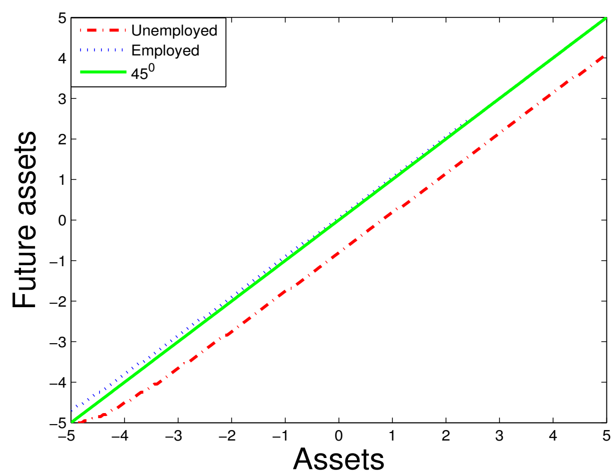


Figure 2.3: **Probabilities associated with optimal effort**

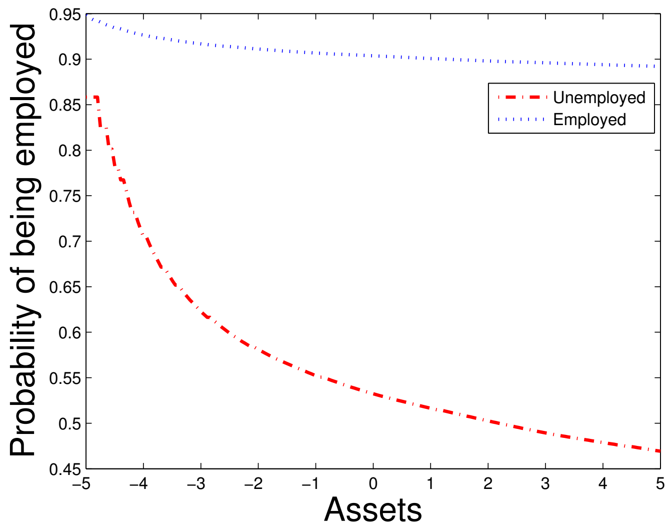


Figure 2.4: **Optimal policy for consumption**

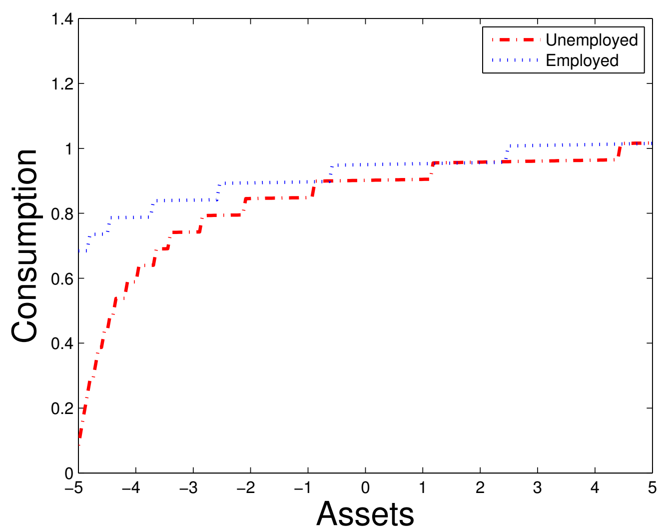


Figure 2.5: **Excess Demand for assets**

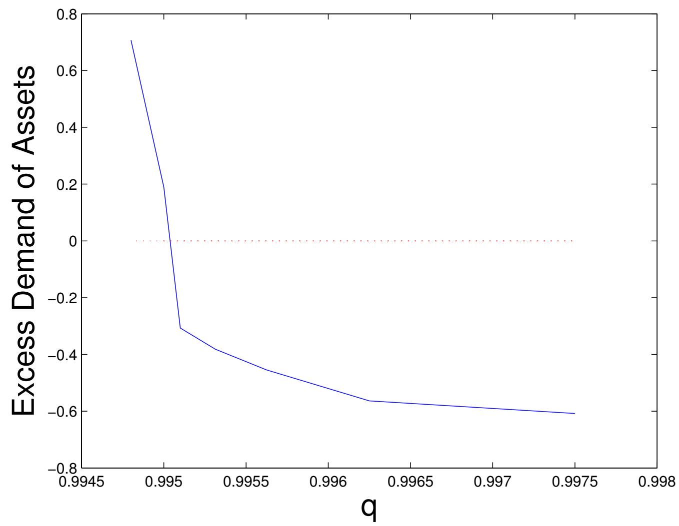


Figure 2.6: **Stationary distribution of assets**

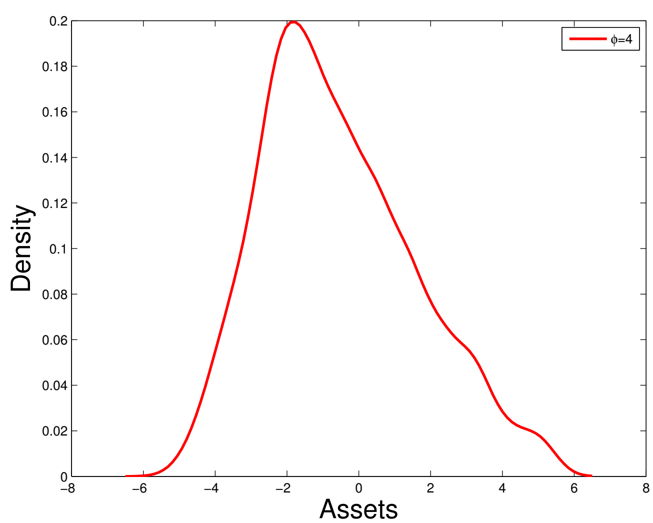
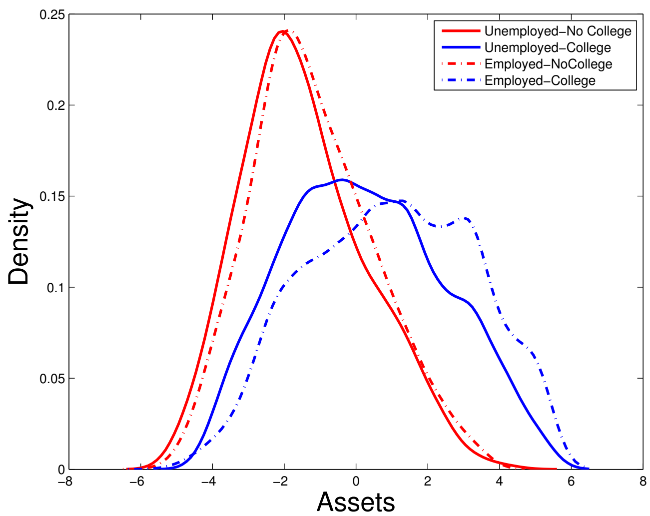


Figure 2.7: **Conditional stationary distribution of assets**



CHAPTER 3

Nonparametric learning when unknown Markov transitions are endogenous

3.1 Introduction

Stochastic dynamic models study environments where agents take actions that affect current payoffs and future states, while maintaining the uncertain nature of the outcomes. Most of these models, usually known as the rational expectations ones, assume individuals know the relevant functions and parameters necessary to take the optimal decision. Although, this is usually not true for most of applications in economics, its motivation depends on the idea that agents have faced the environment enough times so that they have learned all the relevant aspects of it, including the optimal decisions.

However, it has been proven this is not necessarily the case. Two sources for the lack of convergence have been identified. The first of them shows that complete learning may not be optimal. This is the case of the bandit problems studied by [27] when arms (actions) are countable, and later extended by [22] to a more general scenario. The second of them was recently addressed by [1]. They showed that when individuals have an arbitrarily small uncertainty on the conditional distribution of the signals, the difference in the implied asymptotic beliefs is not necessarily small.

In this paper we focus on models where actions determine the Markov transitions over future states but this transitions are unknown. For example, investment and learning are usually simultaneous and interdependent since the returns of different projects are unknown and the only way to learn about them is by undertaking them. This problem is associated with the adoption of new technologies and our approach explains why better technologies are not always implemented in developing countries. Another example that fits this situation is related to the health status of an individual. Health is determined by preventive care when individuals are healthy and by taken medicines when individuals are ill, although their particular effect is uncertain.

The purpose of the paper is to provide a general learning process for this unknown endogenous transitions. We generalize previous results by not imposing parametric restrictions on the distribution over future states. Therefore, our approach is robust to the second source of lack of convergence.

We first consider the case where the possible set of actions is discrete as in bandit problems. We develop a tractable bayesian model where we impose no structure on the transition functions, although it guarantees convergence to the true transition function provided that all actions are taken enough times. Unfortunately, this latter condition is not always satisfied; in our model optimal incomplete learning is also possible.

We then allow the actions to take a convex and compact set of values as is usual in most of the dynamic literature. We enrich the model by giving more structure to the transition function. In particular, we let it be a function of the action taken, a constant initial attribute, and stochastic shocks. The second component captures inherited characteristics, such as ability or genes, that are unknown to individuals and must discover through time. The third component introduces the stochastic nature of the problem. We assume the stochastic shocks are not observable and their density is unknown, as opposed to the traditional literature where a particular density is assumed (a normal one most of the times) and a bayesian learning rule is specified over the parameters that characterize the density.

Our approach is different to what has been done previously since it involves learning about functions and not only parameters.[1] For instance, we do not need assumptions over the density of the shocks that hit the individual each period, yet it plays an enormous role in determining the optimal decisions and can be (partially) learned from our process. To find a learning process we use tools from nonparametric econometrics in a similar way parametric OLS regressions have been used in adaptive learning. This approach guarantees convergence to the rational expectations actions but in a more general framework.

We show that, even though the agents cannot learn all of the unknowns, they will be able to learn the probability of having the good state, which is a composite function of the unknowns and is the relevant function to take the optimal decision. The unknowns are not identified since the relevant probability can be obtained by different compositions of functions. Therefore, we impose some structure on the transformation function of the actions and use a normalization by focusing on its least concave representative, a function that can be estimated by nonparametric techniques [49]. This normalization allows us to find also an estimate of the density of the shocks and the inherited attribute that matches with the overall composition.

Finally, we do not require completely time consistent individuals. Our learning process

[1]One remarkable exception is [17]. However, as discussed later in this section, their approach cannot be applied in our framework.

apply for passive learners who think each period they know the relevant functions to take an optimal decision, although at the beginning of each period they update those functions with the experience they have accumulated. A completely forward looking agent would be able to learn faster and completely by also using the actions for experimentation.

3.1.1 Related literature

Our model is related with literature on bandit problems where agents take sequential actions to maximize the discounted present value of an stream of rewards, but without knowing ex ante the rewards from each action. The tradeoff faced by agents is between exploitation and exploration (or experimentation), which can be associated between short term and long term rewards, respectively. Our model generalizes previous ones by not assuming a particular conditional density over the signals.

As mentioned before, this paper is also related to adaptive learning. In the context of macroeconomics and general equilibrium, the main body of this literature assume agents perform OLS regressions to estimate unknown parameters given the available information these agents have.[2] Several papers have shown how a rational expectations equilibrium can be achieved using this approximation even if agents mispecify the relevant model; see for example [7], [12], or [13]. Furthermore, [46,47] provide conditions when model mispecification is not crucial to the achievement of the equilibrium in an environment where the unknown parameters face linear dynamics.

However, as [40] shows, the convergence to a rational expectations equilibrium is sensitive to the stochastic law of motion of the unknown parameters and the specification of the agent's model that they estimate. They show that if the dynamics of the parameter is non-linear,

[2]There are alternative methods such as constant gain learning. See [24] for a extensive review of these approaches in macroeconomics.

mispecification will lead to non-optimal and non-rational equilibria. This result motivated the introduction of a non-parametric adaptive learning process by [17] that applies both for rational expectations and Nash equilibria. Their paper follows a similar structure of the previous literature but allows the agent to learn the relevant functions to achieve an equilibrium. Although we share the same spirit of nonparametric learning processes, our methods differ. This difference arises because they only consider learning algorithms that evolve linearly.

In our model learning is achieved by estimating in each period the normalized functions using modified maximum likelihood functions in two stages (see [18] and [50]). We follow importantly [49] since her paper deals with the nonparametric estimation and identification of several type of models related to ours. Moreover, she provides the consistency result of the estimator that is necessary to achieve the learning outcome of our paper. Therefore, we use her results as tools to generate learning in our model. However, our model, its implications, and its interpretation are new in the literature.

The organization of the paper is as follows. In the second section we describe the model with discrete actions and nonparametric learning. In the next section we describe the case where actions can take a convex and compact set of values and we introduce the assumptions required to obtain the convergence in this scenario. Finally, in the last section, we conclude.

3.2 Discrete Actions

Consider an infinitely lived agent who wants to maximize her expected utility function over her lifetime. At the beginning of each discrete period a state $s_t \in S$ is observed and the agent chooses an action $a_t \in A$. Let A and S be discrete sets indexed by $n = 1, ..., N$ and $k = 1, ..., K$, respectively. The payoff per period is given by $u\left(a_t; s_t\right)$ and actions also determine the probability distribution over future states. Let $P_n^k(s_j)$ be the probability that the future state is s_j given the current action a_n and current state s_k. Therefore there are $K \times N \times K - 1$ transition probabilities.

Suppose first that the transition probabilities are known. For simplicity of the exposition, let us assume an individual maximizes a Von Neuman- Morgenstern utility function represented recursively by

$$v(s_k) = \max_n \left\{ u(a_n, s_k) + \gamma \sum_{s' \in S} v(s') P_n^k(s') \right\} \text{ for} \gamma \in (0,1)$$

When the distribution is not known we have to keep track of the agent's beliefs. Suppose first there are two states $s_H > s_L$ and two actions, then we just need to keep track of four probabilities. We let $P_n^k(s_H)$ have as prior a Beta distribution with parameters $\alpha_n^k(s_L)$ and $\alpha_n^k(s_H)$. This distribution is defined on the interval $[0,1]$ and its mean is given by

$$\mathbb{E}\left[P_n^k(s_H)\right] = \frac{\alpha_n^k(s_H)}{\alpha_n^k(s_H) + \alpha_n^k(s_L)}$$

The prior is updated using Baye's rule every time a_n is chosen when the current state is s_k and the new s_{t+1} is drawn, if not it will remain the same. This updating rule follows the literature in bandit problems when arms are independent. Given a prior with parameters $\left(\alpha_{n,0}^k(s_L), \alpha_{n,0}^k(s_H)\right)$, the updating rule for $\boldsymbol{\alpha} = \left(\alpha_1^1(S), \alpha_2^1(S), \alpha_1^2(S), \alpha_2^2(S)\right)$ is given by:

$$\alpha_{n,t+1}^k(s_j) = \begin{cases} \alpha_{n,t}^k(s_j) + 1 & \text{if } s_{t+1} = s_j,\ a_t = a_n,\ s_t = s_k \\ \alpha_{n,t}^k(s_j) & \text{otherwise} \end{cases}$$

Note that the state next period plays the role of a signal. The parameter $\alpha_{n,0}^k(s)$ is the initial belief of the number of periods the individual has had state s after choosing action

a_n when the previous state was s_k. When the initial prior is given by $\alpha_{n,0}^k = 1$ and $\beta_{n,0}^k = 1$ it represents a uniform distribution over $[0,1]$, the most uninformative prior. Moreover, conditional on having a previous state s_k and choosing action a_n, the mean $\mathbb{E}\left[P_{n,t}^k\left(s_H\right)\right]$ can be interpreted as the number of times and individual has had a good state plus her initial belief, over the total number of periods (including the beliefs ones).

The generalization of this framework to K states and N actions follows the same structure. Let $P_n^k\left(S\right)$ have as prior a Dirichlet distribution $F\left(\cdot|\alpha_n^k\left(S\right)\right)$ with parameters $\alpha_n^k\left(S\right)$.[3] In this case

$$\mathbb{E}\left[P_n^k\left(s_j\right)\right] = \frac{\alpha_n^k\left(s_j\right)}{\sum_{i=1}^K \alpha_n^k\left(s_i\right)}$$

The updating rule now applies over $\boldsymbol{\alpha} = \left(\alpha_1^1\left(S\right), ..., \alpha_n^k\left(S\right), ..., \alpha_N^K\left(S\right)\right)$ and is the same as before. For convenience we will also denote an updated $\boldsymbol{\alpha}$ as $\boldsymbol{\alpha}_{n,j}^{k\prime}$ when an update has been done to $\alpha_n^k\left(s_j\right)$ whereas the other parameters remain the same. That is, when current state was s_k, action a_n was chosen, and the next period realization was s_j.

The problem can be expressed recursively as

$$\begin{aligned}
v\left(s_k, \boldsymbol{\alpha}\right) &= \max_n \left\{u\left(a_n, s_k\right) + \gamma \mathbb{E}\left[v\left(s_j, \boldsymbol{\alpha}_{n,j}^{k\prime}\right)|a_n, s_k\right]\right\} \\
&= \max_n \left\{u\left(a_n, s_k\right) + \gamma\left[\sum_{i=k}^K \mathbb{E}_{P_n^k(s_j)}\left[P_n^k\left(s_j\right) v\left(s_j, \boldsymbol{\alpha}_{n,j}^{k\prime}\right)\right]\right]\right\} \\
&= \max_n \left\{u\left(a_n, s_k\right) + \gamma \sum_{j=k}^K \left[\frac{\alpha_n^k\left(s_j\right)}{\sum_{i=1}^K \alpha_n^k\left(s_i\right)} v\left(s_j, \boldsymbol{\alpha}_{n,j}^{k\prime}\right)\right]\right\}
\end{aligned}$$

[3]When $K = 2$ the Dirichlet distribution is just a Beta distribution.

Therefore the agent only cares about the frequency each state has been realized given previous actions and states. Note that our learning process did not assume any parametric restriction over the distribution of the signals as a function of the actions. Our parametric assumptions over the prior were only used for each point in the grid, but we did not assume anything about the overall composition. Moreover, our prior assigns positive mass to any number in the interval $(0, 1)$, guaranteeing convergence if the action is taken enough times.

Therefore, this approach is more robust than traditional ones assuming agents know the conditional distribution of the signals. This is particularly important since it has been shown recently by [1] that when individuals do not perfectly identify this conditional distribution and have a small uncertainty about it, asymptotic beliefs can vary importantly from their true value.

However, as in bandit problems, we can not guarantee that the agent will eventually converge to the optimal decision taken in the model where the transitions are known. The intuition behind this result is that ex ante an agent may find optimal to stick with one action. If priors of other actions were not favorable then the individual will not find them attractive to try them enough times so as to be close to the true probabilities. This strategy implies that the individual will learn perfectly the transition associated with the action taken whereas having bad estimates of the transitions associated with other actions. In the next subsection we present an example of non convergence.

3.2.1 Example

Consider an infinitely lived consumer who has to invest in her health every period. Being healthy generates an income $s_H = 1$, whereas an unhealthy person gets half of this income $s_L = 0.5$. This income is used for consumption and investment $c_t + i_t = s_t$. The agent orders the consumption streams according to $U(c) = \sum_{t=1}^{\infty} \beta^t \mathbb{E}_{s_t}\left[\frac{c_t^{1-\sigma}}{1-\sigma}\right]$, where $\sigma = 2$; whereas investment increases the probability of being healthy next period. Therefore, the current payoff is given by $u(i; s) = \frac{(s-i)^{1-\sigma}}{1-\sigma}$.

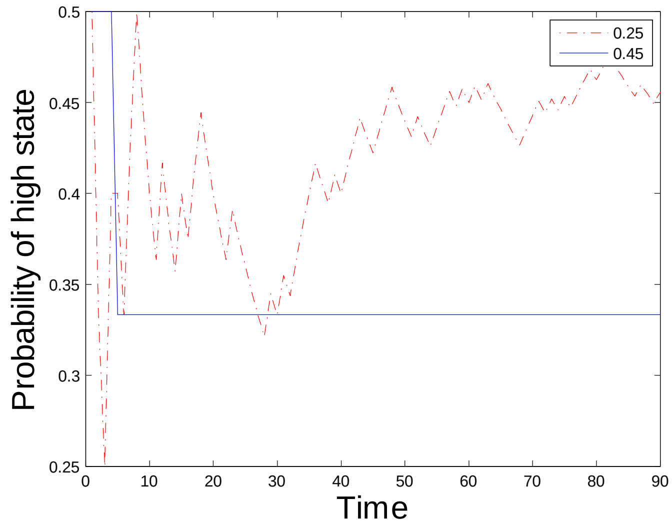
Investment can only take two values: $i_1 = 0.25$ and $i_2 = 0.45$. The probability of being healthy next period conditional on the investment taken is independent of the current state and is given by $P(s_H|i) = 2i$. Let the discount factor be $\beta = 0.9932$. Under this setup, rational expectations optimal decisions consist on choosing the high investment when she has the high state, and choosing the low investment when having the low state.

However, when transitions have to be learned, convergence to the rational expectations decisions is not obtained. We simulate shocks for 100 periods and analize the behavior of this agent in the learning environment described before. Starting with the most uninformative prior for both levels of investment, individuals will first try the cheapest investment since the expected probability is the same (0.5). After getting bad shocks in the next two periods the individual updates this transition and lower its mean (0.25).

The next period the individual tries the high investment but after obtaining a bad shock (which happens with a 0.1 probability) the individual updates the transition probability and lowers its mean to 1/3. After period 3 the individual does not choose the high investment again since its relatively more costly than the low investment. Moreover, the transition associated with the low investment starts to converge and the agents gets stuck with low investment. The next figure shows the simulated evolution of the probabilities.

One can also include nonparametric restrictions such as monotonicity or concavity of the transition function with respect to the actions to refine the priors. In this case actions will be correlated now, which means that an action could be used to learn about the effect of another action. Nevertheless, [22] provide an analysis of settings where incomplete learning

Figure 3.1: **Evolution of beliefs**



may be ex ante optimal even when arms are correlated.

To break this negative result we need a model with more structure that guarantees complete learning on the optimal action. In the next section we explore a model where we impose additional assumptions on the transition function and the set of possible values the actions to ensure the convergence to the optimal action, while maintaining the nonparametric approach.

3.3 Convex and Compact set of actions

Let us now assume the set A is convex and compact. In particular let $A = \prod_{l=1}^{L} [\underline{a}_l, \bar{a}_l] \subset \mathbb{R}^L$ be an L-dimensional closed interval containing the vector of zeros.[4] Let also $u(a_t; s_t)$ be a decreasing and differentiable convex function in a_t, which could be interpreted as a cost function that wants to be minimized. We will now use nonparametric econometrics to keep

[4]The necessity of a being a vector will become clear later.

track of the unknown transition function. For this purpose we focus on a two state model where $s_H > s_L$, we assume the transition between states is independent on the previous state, and we add concavity and monotonicity as nonparametric conditions over the transition as a function of the actions.

The probability of having a high state next period is assumed to be increasing in the effect of an action taken in the current period $h(a_t)$. This function captures the possibly unknown effect of the actions and its mapping to the space of the stochastic shocks. For example, the effect of preventive care may be different between different agents. The function $h : A \to \mathbb{R}^+$ belongs to Ω, the set of concave, monotone increasing, strictly positive and twice continuous differentiable functions satisfying $h(0) = 0$. That is, the set of production functions satisfying non-increasing returns to scale, *no free lunch* and free disposal.[5]

The probability also depends on an inherited and constant through time characteristic of the individual $\theta \in \Theta \subseteq \mathbb{R}$. For example, this parameter can be associated with genes in the case when the state is health. Finally, the state will also depend on i.i.d. shocks $\varepsilon_t \in \mathbb{R}$, which density $F_\varepsilon(\cdot)$ belongs to $\Gamma = \{F : \mathbb{R} \to [0,1] \mid F \text{ is monotone increasing and } F(0) = 0.5\}$.

In particular, suppose the following law of motion for s

$$\begin{aligned} s_{t+1} &= s_H \quad \text{if} \quad \theta + h(a_t) \geq \varepsilon_t \\ s_{t+1} &= s_L \quad \text{otherwise} \end{aligned} \tag{3.1}$$

[5]We do not impose Inada conditions since we do not want to rule out 0 as an optimal investment. In fact, we require that the first derivative of this function has to be bounded.

The additivity between the effect of the actions and the initial attribute is necessary to interpret and partially identify θ after the normalizing assumptions over F_ε and h. Hence, it can be interpreted either as the effect on the probability even when no actions have been taken, or as the mean of the distribution of the shocks. The first interpretation is closer to the case of genes or ability, whereas the second is related to how lucky an individual is or as a characteristic of the environment she faces. We stick with the first one since it goes along the traditional economic literature. Also note that if θ were to enter multiplicatively, for example, then it would be just part of the function h.

Given this law of motion, the probability of having the high state in $t+1$ is given by $F_\varepsilon\left(\theta + h\left(a_t\right)\right)$. We require this probability to be concave in a_t in order to reflect the fact that the marginal probability of the actions is positive but decreasing. This probability also shows how the likelihood of being high type tomorrow depends on initial characteristics and possible shocks we can experience (nature) and the decisions we have taken in our life (nurture).

This is a strictly concave problem satisfying monotonicity and discounting, hence the first order conditions are necessary and sufficient.[6] Assuming the agent knows the environment perfectly, including functions h and $F_\varepsilon\left(\cdot\right)$ and her initial characteristic θ, then the optimal decision $a^*\left(s\right)$ is characterized by

$$-D_a u\left(a^*\left(s\right);s\right) = \beta\left[v\left(s_H\right) - v\left(s_L\right)\right] f_\varepsilon\left(\theta + h\left(a^*\left(s\right)\right)\right) D_a h\left(a^*\left(s\right)\right) \tag{3.2}$$

where D_a denotes the gradient of derivatives since a is a vector.

[6]See [59] for details.

Interestingly, the density of the stochastic shocks may generate non-monotonic comparative statics. For instance, if the distribution of the random shocks follows a normal distribution, the action gets its maximum when $\theta = -h\left(a^*\left(s\right)\right)$. If the agent has a very negative initial attribute, the optimal action will be small because it is too costly to increase the likelihood of having a high state tomorrow and overcome the inherited attribute. On the other hand, if the individual has a very high θ, it is not necessary to take greater actions to insure a high probability of having the good state. In other words, for agents with extreme genes their future state is almost predetermined (decided basically by nature), but for agents with mild genes their state will be importantly determined by nurture. A similar behavior is observed for single peaked distributions. However, if the distribution is uniform, the optimal action is independent of θ.

3.3.1 Identification and Least Concave Representatives

Now assume the agent does not know her genes θ^*, the action production function $h^*\left(\cdot\right)$ and the distribution of the stochastic shocks $F_\varepsilon^*\left(\cdot\right)$. Following [49] closely, we will show that, although these unknowns can not be identified, we can characterize the set of trios $\left(\theta, h, F_\varepsilon\right)$ that are observationally equivalent to $\left(\theta^*, h^*, F_\varepsilon^*\right)$. Therefore, we can focus on a representative of such set.

We choose the least concave as such representative since is the closest to a linear function and thus is easier to estimate using linear programming. Its estimation follows the same spirit as the revealed preference approach presented in [4] where a piecewise concave function is obtained. Moreover, we know that as the observations becomes "dense" it will converge to the true function [48, 49].

Definition 8 *The trio* $\left(\theta, h, F_\varepsilon\right) \in \left(\Theta, \Omega, \Gamma\right)$ *is observationally equivalent to* $\left(\theta^*, h^*, F_\varepsilon^*\right)$ *if*

$F_\varepsilon(\theta + h(a)) = F_\varepsilon^*(\theta^* + h^*(a))$ *for all* $a \in A$.

That is, the trios are observationally equivalent if they yield the same probability of having a high state, which is the object we are interested in. A trio is identified within a set if there is no other trio in that set that generates the same probability. The next lemma shows that this is not the case.

Lemma 9 *There exists* $(\theta, h, F_\varepsilon) \in (\Theta, \Omega, \Gamma)$ *that is observationally equivalent to* $(\theta^*, h^*, F_\varepsilon^*)$ *and such that* $(\theta, h, F_\varepsilon) \neq (\theta^*, h^*, F_\varepsilon^*)$.

Proof. Let $(\theta, h) \in (\Theta, \Omega)$ be such that for some strictly increasing superadditive function $f : \mathbb{R} \to \mathbb{R}$ satisfying $f(0) = 0$, $\theta + h(a) = f(\theta^* + h^*(a))$ and $(\theta, h) \neq (\theta^*, h^*)$. Let $F_\varepsilon = F_\varepsilon^* \circ f^{-1}$, then $(\theta, h, F_\varepsilon) \in (\Theta, \Omega, \Gamma)$ for all $a \in A$. Hence $F_\varepsilon(\theta + h(a)) = F_\varepsilon^*(\theta^* + h^*(a))$ for all $a \in A$, i.e $(\theta, h, F_\varepsilon)$ is observationally equivalent to $(\theta^*, h^*, F_\varepsilon^*)$. ■

Usually authors impose more restrictions over the functions in order to identify them.[7] However, we can characterize all the observationally equivalent trios as the next Lemma suggests.

Lemma 10 *If* $(\theta, h, F_\varepsilon) \in (\Theta, \Omega, \Gamma)$ *is observationally equivalent to* $(\theta^*, h^*, F_\varepsilon^*)$ *then there exists a strictly increasing function* $f : R \to R$ *such that* $f(0) = 0$, $\theta + h(a) = f(\theta^* + h^*(a))$ *and* $(\theta, h) \neq (\theta^*, h^*)$.

Proof. Suppose there is no such increasing f. Then for some $a_1, a_2 \in A$ either $h(a_1) > h(a_2)$ and $h^*(a_1) \leq h^*(a_2)$, or $h(a_1) \leq h(a_2)$ and $h^*(a_1) > h^*(a_2)$. However, since h and h^* are

[7] For a complete survey of restrictions used to achieve identification see [51].

strictly increasing, neither statement can be true, hence such increasing f must exist. In order to have observationally equivalence it must be true that $F_\varepsilon = F_\varepsilon^* \circ f^{-1}$. Hence, we need $f(0) = 0$ in order to have $F_\varepsilon(0) = 0$ and $F_\varepsilon \in \Gamma$. ■

Therefore, we can take one representative of each equivalent class and focus our attention in it without loss of generalization. We choose the least concave function as such representative since there are non parametric econometric techniques to calculate this function. To explain the definition of a least concave function we follow Debreu (1976).

Let U be the set of all monotone transformations of $h : A \to R^+$, a concave function. The least concave function $\bar{h}$ can be defined as the least element in U with respect to the relation "more concave than", where u is more concave than v if there exists a concave function $g : \mathbb{R}^+ \to \mathbb{R}^+$ such that $u = g \circ v$. Formally:

Definition 11 *Suppose $h : A \to \mathbb{R}^+$ is a concave function. A function $\bar{h} : A \to \mathbb{R}^+$ is a least concave representative if for any strictly increasing function $f : \mathbb{R}^+ \to \mathbb{R}^+$ such that $f \circ h$ is concave, there exists a concave function $g : R^+ \to R^+$ such that $f \circ h = g \circ \bar{h}$.*

This definition implies that if a concave function is a strictly increasing transformation of other, they share the same least concave representatives. [19] also showed the existence, uniqueness up to increasing linear transformation and continuous differentiability when h is continuous differentiable of the least concave representatives.

Note that the definition of least concave representatives is closely related to the environment described in this paper considering that the probability of having high state is a composite function of an increasing function F_ε and a concave function $\theta + h(a)$. Also note that since we impose the restriction that $h(0) = 0$, the least concave representative is unique up to a constant (not a linear transformation).

In order to obtain a nonparametric estimator of the least concave function we need to impose compactness in all the sets we are interested in. We first define a compact set $W \subset \Omega$ of twice continuous differentiable least concave functions. To obtain it we impose some plausible boundary and Lipschitz conditions on its values, gradients, Hessians and Gaussian curvatures (see [49] for details). Also let $\theta \in [\theta_{\min}, \theta_{\max}] = \Theta$ and $\hat{\Gamma}$ be the restriction of Γ to the compact domain. Finally, functions in W must satisfy $\theta_{\min} + h(\underline{a}) = \underline{x}$ and $\theta_{\max} + h(\bar{a}) = \bar{x}$, where $\underline{x}$ and $\bar{x}$ are known.

Let $\hat{W}$ be the set of concave functions that possess least concave representatives in W. Therefore, all functions in $\hat{W}$ possess a unique representative in W. The next lemma provides a useful characterization of observationally equivalent concave functions.

Lemma 12 *Suppose that* $(\theta, h, F_\varepsilon) \in \left(\Theta, \hat{W}, \hat{\Gamma}\right)$. *Let* $\bar{h}$ *and* $\bar{h}^*$ *denote respectively the least concave representatives of* h *and* h^*. *Define* $\bar{F}_\varepsilon$, $\bar{F}_\varepsilon^*$, $\bar{\theta}$ *and* $\bar{\theta}^*$ *by* $\bar{F}_\varepsilon\left(\bar{\theta} + \bar{h}(a)\right) = F_\varepsilon(\theta + h(a))$ *and* $\bar{F}_\varepsilon^*\left(\bar{\theta}^* + \bar{h}^*(a)\right) = F_\varepsilon(\theta^* + h^*(a))$. *Then* $(\theta, h, F_\varepsilon)$ *is observationally equivalent to* $(\theta^*, h^*, F_\varepsilon^*)$ *if and only if* $\left(\bar{\theta}, \bar{h}, \bar{F}_\varepsilon\right) = \left(\bar{\theta}^*, \bar{h}^*, \bar{F}_\varepsilon^*\right)$.

Proof. Suppose $(\theta, h, F_\varepsilon)$ is observationally equivalent to $(\theta^*, h^*, F_\varepsilon^*)$. By Lemma 10, there exists a strictly increasing function $f : \mathbb{R} \to \mathbb{R}$ such that $f(0) = 0$ and $\theta + h(a) = f(\theta^* + h^*(a))$. By Definition 11, the uniqueness of the least concave representative and the definition of W we have that $\bar{\theta} = \bar{\theta}^*$ and $\bar{h} = \bar{h}^*$. Finally, by definition of $\bar{F}$ and $\bar{F}^*$, and observationally equivalence between the trios, it is true that $\bar{F}^*\left(\bar{\theta}^* + \bar{h}^*(a)\right) = \bar{F}\left(\bar{\theta} + \bar{h}(a)\right)$. Hence $\bar{F} = \bar{F}^*$.

Conversely, suppose $\left(\bar{\theta}, \bar{h}, \bar{F}_{\varepsilon}\right) = \left(\bar{\theta}^{*}, \bar{h}^{*}, \bar{F}_{\varepsilon}^{*}\right)$. Then, by definition of $\bar{F}_{\varepsilon}$ and $\bar{F}_{\varepsilon}^{*}$, $F_{\varepsilon}\left(\theta^{*} + h^{*}\left(a\right)\right) = \bar{F}_{\varepsilon}^{*}\left(\bar{\theta}^{*} + \bar{h}^{*}\left(a\right)\right) = \bar{F}_{\varepsilon}\left(\bar{\theta} + \bar{h}\left(a\right)\right) = F_{\varepsilon}\left(\theta + h\left(a\right)\right)$, implying the observationally equivalence of $\left(\theta, h, F_{\varepsilon}\right)$ and $\left(\theta^{*}, h^{*}, F_{\varepsilon}^{*}\right)$. ■

This Lemma implies that all observationally equivalent trios giving us the same probability of having the high state have the same least concave representative. Moreover, the genes can be compared across individuals if we calculate that least concave representative. Also, since $\bar{h}^{*}$ is continuously differentiable, the marginal probability of the actions using the least concave representatives is the same as the one using the true functions. Thus, the agent can calculate Equation (3.2) accurately with this information.

Let $y_t = 1$ if $s_{t+1} = s_H$ and 0 if $s_{t+1} = s_L$. [49] shows that we can obtain a strongly consistent and identifiable estimator of $\left(\bar{\theta}^{*}, \bar{h}^{*}, \bar{F}_{\varepsilon}^{*}\right)$ with the observables $\left\{\left(y_t, a_t\right)\right\}_{t=1}^{T}$ if the following assumptions are satisfied:

1. The random term ε is independent of a has a strictly increasing cumulative distribution F_{ε}^{*}.
2. Actions a possesses a bounded Lebesgue density g whose support is A and whose probability measure P_X is absolutely continuous.
3. The function h^{*} belongs to the set of all concave functions that posses least concave representatives in W.

All these properties are satisfied in our model by construction. The estimation procedure for the least concave function and the initial attribute uses the maximum score estimation method, whereas the density of the shocks is obtained via maximum likelihood (see [49] for details). Throughout this section we have assumed a is a vector. If a where to belong to the real line, the least concave function is simply given by a linear function connecting the two known limit values of the function $\theta_{\min} + h(\underline{a})$ and $\theta_{\max} + h(\bar{a})$, which will also allow us to obtain θ^*. Therefore, an estimation of the least concave function is nontrivial only if a is a vector.

3.3.2 Passive Learning

Now suppose the agent does not know the trio $(\theta^*, h^*, F_\varepsilon^*)$. We focus on passive learners which we define as agents who do not internalize the learning process and thus does not experiment and learn optimally. This agent takes the decision as if she knows the true functions. After more observations are obtained she updates her beliefs but again thinks those are the true unknown functions.

We follow this approach since the internalization of the learning process involves the inclusion of a infinite dimensional object as a state variable. Despite this inconsistency, this type of agents will be arbitrarily close to the optimal decision as she accumulates experience in this scenario. Consistent or active learners will be able to completely learn at a greater speed by doing experimentation.

We have shown that, although we cannot identify all the unknown functions and parameters, we can focus on a representative of the class of functions that yield the same probability, that is we will be only interested in $F^*(a) = F_\varepsilon^*(\theta^* + h^*(a))$. Then, by generating observations through time via the choices made by the individual and the experimentation made by nature, the agent will be able to estimate consistently the least concave representative. Therefore, as long as the individual does not get stuck in an action, she would be able to approach the optimal decision as time goes to infinity. The reason why the agents do not get stuck is because of the smoothness properties imposed to the least concave function, which are inherited to the transition function.[8]

In time zero, let the agent have a prior over $(\theta, h, F_\varepsilon) \in \left(\Theta, \hat{W}, \hat{\Gamma}\right)$. [9] These priors can be interpreted as information learned from others experiences. Using these priors and according to the first realization $s_0 = s_k$ the individual calculates the optimal action $a_0(s_k)$ using Equation (3.2). She will do the same calculation every period until her state changes, say time τ. Then $s_\tau = s_{-k}$ and her optimal action $a_\tau(s_{-k})$ changes, where "$-k$" denotes the opposite state of s_k. Of course we need the priors to imply different optimal actions in each state to have a meaningful learning process. At time $\tau + 1$ the agent will have enough variation to estimate a new prior of the functions using her history $\{s_{t+1}, a_t\}_{t=0}^{\tau}$ and the technique described in [49].

With the new updated prior, she will use Equation (3.2) to calculate the current optimal action $a_{\tau+1}$. After that period, she will update her prior every period t with the new information she gathers, (s_{t+1}, a_t). The smoothness properties inherited by the transition function will generate a contraction and will imply the convergence of the action to the optimal one. However, since optimal decision will tend to concentrate around $(a^*(s_L), a^*(s_H))$ as the individual approaches to the true functions, the estimation will be dominated by these observations and therefore will ruin its properties.

To avoid this result we can define an optimal stopping time that allows us to be as close as we want to the true functions. Formally, for any $\varepsilon > 0$ there exists a T such that $d(\digamma_T, \digamma^*) < \varepsilon$. After T the individual will stop learning and will conform to be sufficiently close to the optimal actions. Note this is a consequence of this passive learning process

[8]Of course the utility function has to be smooth enough. In this case, concavity and differentiability will suffice.

[9]Note the prior we are defining here does not follow the bayesian sense. In the bayesian approach, a prior should be defined as a nondegenerate probability distribution over the sets. In this case is just a degenerate probability distribution over the sets.

3.3.3 Optimal Learning

This learning process is myopic since the agent does not internalize the updating process generated every period by the estimation. An active learner would consider this updating process as a constraint every period and the estimated functions and parameters as state variables. This refinement would allow for experimentation since the F.O.C. (Equation (3.2)) used in the optimum under complete knowledge gets distorted. In this case the choice of the action will also consider its effect on the updating process. Therefore, the individual will be able to learn faster when learning is internalized.

Formally, the value function for an active learner is given by:

$$v\left(s,\digamma\right)=\max_{a\in A}\left\{\begin{array}{c} u\left(a;s\right)+ \\ \beta\left[F_{\varepsilon}\left(\theta+h\left(a\right)\right)v\left(s_H,\digamma_H\right)+\left(1-F_{\varepsilon}\left(\theta+h\left(a\right)\right)\right)v\left(s_L,\digamma_L\right)\right] \end{array}\right\}$$

subject to $\digamma_k=G\left(\digamma,s_k,a\right)$ for $k=L,H$, the updating process. The optimal decision is characterized now by following F.O.C

$$-D_a u\left(a^*;s\right)=\beta\left\{\begin{array}{c} \left[v\left(s_H,\digamma_H\right)-v\left(s_L,\digamma_L\right)\right] \\ \left[f_{\varepsilon}\left(\theta+h\left(a^*\right)\right)D_a h\left(a^*\right)+E_{s'}\left[D_a G\left(\digamma_{s'},s',a^*\right)\right]\right] \end{array}\right\}$$

Now the Euler equation gets distorted by the term $E_{s'}\left[D_a G\left(\digamma,s',a^*\right)\right]$, which shows the expected benefit from experimentation. This distortion carries the optimal decision away from the rational expectations decision $\left(a^*\left(s_L\right),a^*\left(s_H\right)\right)$. Therefore the properties of the estimator will not be ruined and the agent will be able to learn completely the unknowns.

Nevertheless, as previously shown, the difference of the long run average payoff between the passive and the active learner will be just a function of ε. On the other hand, modelling optimal learning gets tedious since an infinite dimensional variable has to be added as a state, and we can not assure the new constraint have desired properties (such as concavity).

3.4 Concluding remarks

We have studied a dynamic stochastic model where current actions determine the probability distribution over future states, while constrained by a current state . We first studied a bandit problem with possibly correlated arms and nonparametric learning. That is, we did not assume a particular functional of the transition function; instead, we show how to update each point of the transition function using a simple tractable model. This approach has the advantage to be robust to uncertainty on the conditional distribution of the signals, which has been proved to be determinant in the convergence of the asymptotic beliefs [1].

As the classic result in bandit problems, convergence to the optimal action when transitions are known is not obtained. Therefore, we enrich our model by allowing the actions to be in a compact and convex set and by imposing regularity conditions on the transition function. We focus on a two state model and let the probability of having a high type next period depend on the effect of the action chosen by the agent (nurture), a constant inherited attribute and the distribution of the stochastic shocks determined by nature.

We suggested a learning process when the agent does not any of the three functions and the individual does not internalize the learning process. Since these unknowns can not be identified, we first characterized the set of all the unknowns that are observationally equivalent, i.e, that yield the same probability. Secondly, we chose a representative of such set, the least concave one, and show how we can estimate it provided some plausible conditions.

These estimators are consistent, meaning that as the agent gets more experience in this game she can better calculate the representatives. Finally we stressed the difference when the individual internalizes the updating process.

References

[1] D. Acemoglu, V. Chernozhukov, and M. Yildiz. Fragility of asymptotic agreement under bayesian learning. *Working Paper*, 2009.

[2] D. Acemoglu and R. Shimer. Productivity gains from unemployment insurance. *European Economic Review*, 44(7):1195–1224, 2000.

[3] J.T. Addison and P. Portugal. Job displacement, relative wage changes, and duration of unemployment. *Journal of Labor Economics*, 7(3):281–302, 1989.

[4] S.N. Afriat. The construction of utility functions from expenditure data. *International Economic Review*, 8(1):67–77, 1967.

[5] S.R. Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659, 1994.

[6] Y. Algan, A. Chéron, J.O. Hairault, and F. Langot. Wealth effect on labor market transitions. *Review of Economic Dynamics*, 6(1):156–178, 2003.

[7] R.M. Anderson and H. Sonnenschein. Rational expectations equilibrium with econometric models. *The Review of Economic Studies*, 52(3):359–369, 1985.

[8] S. Athey. Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1):187, 2002.

[9] H.G. Bloemen. The relation between wealth and labour market transitions: an empirical study for the netherlands. *Journal of Applied Econometrics*, 17(3):249–268, 2002.

[10] H.G. Bloemen and E.G.F. Stancanelli. Individual wealth, reservation wages, and transitions into employment. *Journal of Labor Economics*, 19(2):400–439, 2001.

[11] A. Bonatti and J. Horner. Collaborating. *The American Economic Review*, 101(2):632–663, 2011.

[12] M. Bray. Learning, estimation, and the stability of rational expectations. *Journal of economic theory*, 26(2):318–339, 1982.

[13] M.M. Bray and N.E. Savin. Rational expectations equilibria, learning, and model specification. *Econometrica: Journal of the Econometric Society*, pages 1129–1160, 1986.

[14] D. Card and D. Sullivan. Estimating the effect of subsidized training on movements in and out of employment. *Econometrica*, 56(3):497–530, 1988.

[15] H. Chade and N. Kovrijnykh. Information acquisition, moral hazard, and rewarding for bad news. *Working Paper*, 2011.

[16] G. Chamberlain and C.A. Wilson. Optimal Intertemporal Consumption under Uncertainty* 1. *Review of Economic Dynamics*, 3(3):365–395, 2000.

[17] X. Chen and H. White. Nonparametric adaptive learning with feedback. *Journal of Economic Theory*, 82(1):190–222, 1998.

[18] S.R. Cosslett. Distribution-free maximum likelihood estimator of the binary choice model. *Econometrica: Journal of the Econometric Society*, pages 765–782, 1983.

[19] G. Debreu. Least concave utility functions. *Journal of Mathematical Economics*, 3(2):121–129, 1976.

[20] J.S. Demski and D.E.M. Sappington. Delegated expertise. *Journal of Accounting Research*, pages 68–89, 1987.

[21] Z. Dong, C. Wang, and F. Xie. Do executive stock options induce excessive risk taking? *Journal of Banking & Finance*, 34(10):2518–2529, 2010.

[22] D. Easley and N.M. Kiefer. Controlling a stochastic process with unknown parameters. *Econometrica: Journal of the Econometric Society*, pages 1045–1064, 1988.

[23] F. Ederer. *Launching a Thousand Ships: Incentives for Parallel Innovation*. Working Paper, 2008.

[24] G.W. Evans and S. Honkapohja. *Learning and expectations in macroeconomics*. Princeton Univ Pr, 2001.

[25] E. Feess and M. Walzl. Delegated expertise–when are good projects bad news? *Economics Letters*, 82(1):77–82, 2004.

[26] A.D. Foster and M.R. Rosenzweig. Learning by doing and learning from others: Human capital and technical change in agriculture. *The Journal of Political Economy*, 103(6):1176–1209, 1995.

[27] J.C. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 148–177, 1979.

[28] J. Gomes, J. Greenwood, and S. Rebelo. Equilibrium unemployment. *Journal of Monetary Economics*, 48(1):109–152, 2001.

[29] R.M. Gritz. The impact of training on the frequency and duration of employment. *Journal of Econometrics*, 57(1-3):21–51, 1993.

[30] D. Gromb and D. Martimort. Collusion and the organization of delegated expertise. *Journal of Economic Theory*, 137(1):271–299, 2007.

[31] G.D. Hansen and A. İmrohoroğlu. The role of unemployment insurance in an economy with liquidity constraints and moral hazard. *The Journal of Political Economy*, 100(1):118–142, 1992.

[32] J. Heathcote. Fiscal policy with heterogeneous agents and incomplete markets. *Review of Economic Studies*, 72(1):161–188, 2005.

[33] J. Heckman, L. Lochner, and C. Taber. Explaining Rising Wage Inequality: Explanations With A Dynamic General Equilibrium Model of Labor Earnings With Heterogeneous Agents. *Review of Economic Dynamics*, 1(1):1–58, 1998.

[34] H.A. Hopenhayn and J.P. Nicolini. Optimal unemployment insurance. *Journal of political economy*, 105(2):412–438, 1997.

[35] H.A. Hopenhayn and E.C. Prescott. Stochastic monotonicity and stationary distributions for dynamic economies. *Econometrica: Journal of the Econometric Society*, pages 1387–1406, 1992.

[36] M. Huggett. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of economic Dynamics and Control*, 17(5-6):953–969, 1993.

[37] R.D. Innes. Limited liability and incentive contracting with ex-ante action choices* 1. *Journal of economic theory*, 52(1):45–67, 1990.

[38] S. Karlin and H. Rubin. The theory of decision procedures for distributions with monotone likelihood ratio. *The Annals of Mathematical Statistics*, 27(2):272–299, 1956.

[39] P. Krusell and A.A. Smith. Income and wealth heterogeneity in the macroeconomy. *The Journal of Political Economy*, 106(5):867–896, 1998.

[40] C.M. Kuan and H. White. Adaptive learning with nonlinear dynamics driven by dependent processes. *Econometrica: Journal of the Econometric Society*, pages 1087–1114, 1994.

[41] R.A. Lambert. Executive effort and selection of risky projects. *The Rand Journal of Economics*, pages 77–88, 1986.

[42] R. Lentz and T. Tranæs. Job search and savings: Wealth effects and duration dependence. *Journal of Labor Economics*, 23(3), 2005.

[43] L. Ljungqvist and T.J. Sargent. *Recursive macroeconomic theory.* The MIT press, 2004.

[44] J.M. Malcomson. Principal and expert agent. *The BE Journal of Theoretical Economics*, 9(1):17, 2009.

[45] G. Manso. Motivating innovation. *The Journal of Finance*, 66(5):1823–1860, 2011.

[46] A. Marcet and T.J. Sargent. Convergence of least-squares learning in environments with hidden state variables and private information. *The Journal of Political Economy*, pages 1306–1322, 1989.

[47] A. Marcet and T.J. Sargent. Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic theory*, 48(2):337–368, 1989.

[48] A. Mas-Colell. On revealed preference analysis. *The Review of Economic Studies*, 45(1):121–131, 1978.

[49] R.L. Matzkin. *Least concavity and the distribution-free estimation of nonparametric concave functions*. Yale University, Cowles Foundation for Research in Economics, 1990.

[50] R.L. Matzkin. Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica: Journal of the Econometric Society*, pages 239–270, 1992.

[51] R.L. Matzkin. Nonparametric identification. *Handbook of Econometrics*, 6:5307–5368, 2007.

[52] R. Mehra and E.C. Prescott. The equity premium: A puzzle* 1. *Journal of monetary Economics*, 15(2):145–161, 1985.

[53] P.R. Milgrom. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391, 1981.

[54] G. Moscarini and L. Smith. The optimal level of experimentation. *Econometrica*, 69(6):1629–1644, 2001.

[55] D. Neal. Industry-specific human capital: Evidence from displaced workers. *Journal of labor Economics*, pages 653–677, 1995.

[56] N. Persico. Information acquisition in auctions. *Econometrica*, 68(1):135–148, 2000.

[57] S. Rendon. Job search and asset accumulation under borrowing constraints*. *International Economic Review*, 47(1):233–263, 2006.

[58] E.G.F. Stancanelli. Do the rich stay unemployed longer? an empirical study for the uk. *Oxford Bulletin of Economics and Statistics*, 61(3):295–314, 1999.

[59] N.L. Stokey, R.E. Lucas, and E.C. Prescott. *Recursive methods in economic dynamics*. Harvard Univ Pr, 1989.

[60] D. Szalay. Contracts with endogenous information. *Games and Economic Behavior*, 65(2):586–625, 2009.

[61] C. Wang and S. Williamson. Unemployment insurance with moral hazard in a dynamic economy. In *Carnegie-Rochester Conference Series on Public Policy*, volume 44, pages 1–41. Elsevier, 1996.

[62] D. Williams. *Probability with martingales*. Cambridge Univ Pr, 1991.

[63] E. Wolff. Recent trends in household wealth in the united states: Rising debt and the middle-class squeezean update to 2007. *Annandale-on-Hudson, NY: Levy Economics Institute, Bard College*, 2010.