Title
Essays in Identification, Game Theory and Love

Permalink
https://escholarship.org/uc/item/61q75165

Author
de Brito Caruso, TIAGO

Publication Date
2016

Peer reviewed|Thesis/dissertation
Essays on Identification, Game Theory and Love

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Tiago de Brito Caruso

2016
© Copyright by
Tiago de Brito Caruso
2016
This dissertation consists of three essays divided into chapters. In chapter 1, I construct a theoretical model of dating and falling in love in an online environment. A relationship is an experimental, common value good and partners slowly learn its quality by observing private news. They always have the option to unilaterally break-up, go back to the website and, after some friction, be matched again with another partner. I construct the unique symmetric equilibrium, which consists of a honeymoon and a break-up phase. I also show the value of the website in the two extremes scenarios: non-friction dating and arranged marriage. Moreover, this value is non-monotonic in the friction.

In Chapter 2, which is co-authored with Greg Kubitz, we study a binary choice model where an agent makes a decision that is informed by his beliefs after observing a public signal. This model generalizes to a wide range of economic environments which require econometricians to estimate the beliefs of agents. With minimal structure imposed on the agent’s utility function, we characterize the structure of information needed to identify the beliefs of the agent after observing both signals and decisions. We find that the information must be sufficiently convincing and dense for the agent’s beliefs to be point identified. When the full range of information is relaxed, we show how the agents beliefs
can be partially identified. Additionally, we explicitly show how the econometrician can construct the sharpest boundaries around the agent’s beliefs as she observes signals and decisions.

In chapter 3, I go back to the topic of playing with altruism. There I construct a simple model of decision making with altruism and show that the utility of the agent receiving the gift can decrease with his initial endowment. The idea is somewhat straightforward and it becomes especially relevant when the gift is indivisible. When the agent allocating the gift is altruistic and the candidate to receive the good is sufficiently poor, he will get the gift. Knowing the decision rule in the second stage, the agent receiving the gift has incentives to manipulate his outside option in the first period. I show that the amount of manipulation is always positive and increasing with altruism.
The dissertation of Tiago de Brito Caruso is approved.

Sushil Bikhchandani

Simon Adrian Board

Jernej Copic

William R. Zame, Committee Co-chair

Rosa Liliana Matzkin, Committee Co-chair

University of California, Los Angeles

2016
To Diana.
# Table of Contents

1 **Optimal Dating Website** ............................................... 1
   1.1 Introduction ..................................................... 1
   1.2 Model ............................................................ 7
      1.2.1 How Beliefs Evolve ........................................ 8
      1.2.2 The Value of the Website .................................. 9
      1.2.3 Value Functions ............................................ 10
      1.2.4 Decision Problem and Equilibrium Concept .............. 11
   1.3 The Analysis ..................................................... 12
      1.3.1 System of Equations ....................................... 12
      1.3.2 Results ..................................................... 12
      1.3.3 Properties of the Equilibrium Path and Asymptotic Behavior . . . 15
   1.4 Simulations ....................................................... 17
   1.5 Welfare ............................................................ 21
   1.6 Discussion ........................................................ 26
      1.6.1 Asymmetric Equilibrium ................................... 26
      1.6.2 Communication .............................................. 29
      1.6.3 Relationship with other economic institutions ............ 31
   1.7 Mathematical Appendix .......................................... 32

2 **Identification of Beliefs on Decision Making** ....................... 41
   2.1 Introduction ..................................................... 41
2.2 Identification of Beliefs .............................................. 45
  2.2.1 Partial identification without full range ...................... 49
  2.2.2 Partial identification with finite information ............... 50
  2.2.3 Multiple States .................................................. 51
  2.3 Identification of priors ............................................. 54
  2.4 Concluding remarks ................................................. 57

3 Playing with Altruism .................................................. 61
  3.1 Introduction ......................................................... 61
  3.2 Divisible Goods Model .............................................. 64
  3.3 Indivisible Goods Model ........................................... 67
  3.4 Manipulative Agents ............................................... 70
  3.5 Conclusion ......................................................... 73
  3.6 Mathematical Appendix ............................................ 74

Bibliography ............................................................... 76
LIST OF FIGURES

1.1 Value Function of Dating. ................................................. 8
1.2 Value Function of Dating. ................................................. 18
1.3 Value Function of Love. .................................................. 19
1.4 Beliefs in both phases. ................................................... 20
1.5 Hazard Rate of Breaking Up. ............................................ 21
1.6 Value Function of love for passive partner. ......................... 28

2.1 Constructing the bounds on priors. ................................. 56

3.1 Utility functions .......................................................... 68
3.2 Non-monotonicity of $g(d_2)$ .......................................... 70
3.3 Manipulative Agents ...................................................... 71
LIST OF TABLES

1.1 Comparing different Dating Websites. . . . . . . . . . . . . . . . . . . . 2
ACKNOWLEDGMENTS

I am very grateful for my advisor, Professor Rosa Matzkin, for her knowledge, support and patience. I will be forever indebted to Professor Jernej Copic, for always believing and pushing me, and for his endless passion for economics and life.

I also grateful for the guidance and help of my Professor and co-Chair of my committee, Bill Zame, and to Professors Simon Board, Sushil Bikhchandani, Moritz Meyer-ter-Vehn, Joseph Ostroy and Tomasz Sadzik.

The PhD. program became manageable because of the company of my friends Marcus Studart, Kyle Woodward, Amanda Ngyuen, Jesper Sorensen, Kelly Leong and Omer Ali with whom I shared many moments of anxiety, stress and joy.

I am especially thankful to a couple of friends who changed me along these years: Carlos Cantu has a discipline and work ethic that is inspirational; and Greg Kubitz, who carried me on his back a few times, mathematically and physically.

I could not have completed my studies without my parents and siblings who supported me emotionally through my life and without the love and wit that I get everyday from Diana.
2007  B.A. (Economics), PUC-Rio, Rio de Janeiro, Brazil.

2010  L.L.B. (Law), UERJ, Rio de Janeiro, Brazil.

2010  M.A. (Economics), PUC-Rio, Rio de Janeiro, Brazil.
CHAPTER 1

Optimal Dating Website

1.1 Introduction

With whom we choose to spend our lives is perhaps the most important choice we will ever make. Nowadays, the majority of the people looking to make this crucial decision are on dating websites. These websites or mobile apps differ from each other in many dimensions, in this chapter we will focus on one aspect: how frequent should they offer new partners.

Initial economic intuition tells us that the easier it is to find a new partner, the better. However, the most popular websites and apps have very different approaches to this topic. The following table summarizes the distinction that we will explore in this chapter. Match.com allows members to freely scan through their members, whereas E-Harmony, launched five later, constrains your choices to their recommendations, on average eight per day. More recently, the online dating scene moved to dating apps that use geolocation to offer potential partners. Tinder was the one of the pioneers in this method and it gives you as many partners as you can browse. However, Coffee-Meets-Bagel only allows one match per day and Hinge, another app launched in 2014, allows ten\(^1\). I construct a theoretical model to address precisely this issue and to characterize the website that is optimal for the partners\(^2\).

\(^1\)E-Harmony, Hinge, and The League do not have an explicit constraint in the number of matches offered per day, the average number reported here comes from the reviews in the Apple store.

\(^2\)These sites also differ in other relevant dimensions which I abstract away. Hinge only looks for potential partners among the friends of your friends, The League only allows members who come from elite
Table 1.1: Comparing different Dating Websites.

<table>
<thead>
<tr>
<th>Dating Website</th>
<th>Alexa Ranking*</th>
<th>Matches per Day**</th>
<th>Year Foundation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match.com</td>
<td>316</td>
<td>Unlimited</td>
<td>1995</td>
</tr>
<tr>
<td>E-Harmony</td>
<td>1,204</td>
<td>8</td>
<td>2000</td>
</tr>
<tr>
<td>OkCupid</td>
<td>190</td>
<td>Unlimited</td>
<td>2004</td>
</tr>
<tr>
<td>Tinder</td>
<td>13,581</td>
<td>Unlimited</td>
<td>2012</td>
</tr>
<tr>
<td>Coffee-Meets-Bagel</td>
<td>29,170</td>
<td>1</td>
<td>2012</td>
</tr>
<tr>
<td>Hinge</td>
<td>43,976</td>
<td>10</td>
<td>2013</td>
</tr>
<tr>
<td>The League</td>
<td>103,022</td>
<td>5</td>
<td>2014</td>
</tr>
</tbody>
</table>

* Ranking in November 2015, ** Average based on reviews

The main contribution of the chapter is to develop a model of dating and falling in love that helps us understand why these websites should constraint their members choices. As far as I am aware this is the first work to directly address this topic. To do so, I borrow tools from the bandit and optimal experimentation literature and apply them to describe the optimal dating website. In the last section, I discuss how the conclusions of this model can be extrapolated to other economic institutions.

I also identify an interesting economic force. In this model, agents are not allowed to communicate their feelings, so they often have different perceptions about the quality of the match. This restriction is sufficient to push agents to date for a shorter length than what would be socially desirable and this inefficiency leads to my main result, that dating websites that offer fewer choices can be better than those offering more. The model also develops new empirical predictions about the frequency of break-ups and those conditions can be tested with an appropriate data set.

In section 2 I describe the model. I treat relationships as common value experiment schools, Match.com is paid, OkCupid is free, among other
goods. Partners are introduced by the website and start dating, unaware of the quality of the match. Time is continuous. If they are in a good relationship, private good news can, in a bad one, they never come. Once private news arrives, the partner is sure that he is in a good relationship, but he cannot communicate his impression.

Partners can break up unilaterally at any time. This is the only decision they make, and once one of them has decided to break-up, they both go back to the dating website. In the initial analysis, the value of the dating website is exogenous, but later in the chapter, we set it to be equal to the value of dating a new partner, discounted by the time that it takes to find a new match. This discount factor is how we model the amount of friction that the dating website imposes. We interpret the restrictions on profile browsing as proxies for this friction. In our language, E-Harmony and Coffee-Meets-Bagel would impose more friction on its members than Match.com and Tinder.

In section 3 we develop the main analysis of the chapter. Our model can be summarized in 3 equations: the value function of those who are aware that they are in a good relationship, or in love; the value function of those unaware, or simply dating; and Bayes Rule. Only partners not in love would ever break-up. We show that in this game there is no pure strategy symmetric equilibrium, so there is no time $t$ when either partner breaks up for sure. The reason for this result is that if there were such a time, partners considering breaking up would wait a little more just to discover whether the other is in love. This and other features of the model are similar to the analysis of Bar-Isaac [2003].

The game does have a unique symmetric mixed strategy equilibrium. In proposition 1 we describe this equilibrium where partners in love never break-up. Similarly to McAdams [2010], there is a honeymoon phase, when nobody breaks up that lasts until $t^*$. However, after $t^*$, partners not in love are indifferent between breaking-up and keep dating, thus, they randomize their exit decision so that other partners not in love will be also indifferent.

---

3 This a useful but not necessary condition and we will discuss it in section 6
I show that the breaking-up phase never ends, so that partners not in love are always leaving and asymptotically partners in love are becoming more certain that their partners are also in love. This result generates a natural boundary condition that allows us to show the following properties of the equilibrium path in proposition 2: during the break-up phase, the value function of being in love will be increasing, whereas the beliefs and the hazard rate of breaking-up will be decreasing. An external researcher would only be able to observe unconditional break-up rates, so we briefly described this hazard rate path. We also derive the asymptotic behavior of this equilibrium in proposition 3.

In section 4, I choose parameters such that the honeymoon phase lasts for three months and simulate the model. This exercise helps to visualize the equilibrium described in section 3. I show how the value functions, beliefs, and break-up rates evolve over time in both phases.

In Section 5, I analyze the welfare properties of this model. There I treat the value of the dating website as an endogenous variable, it is simply the value of dating at time 0 discounted by the amount of friction that the website imposes. I introduce two extreme benchmarks to help in this analysis: the non-friction matching system, where partners would obtain another match immediately, and the arranged marriage system, where is impossible to obtain another match.

I solve for the value of the website in a non-friction model and in the arranged marriage. If the former benchmark, partners will never find love and will be perpetually breaking-up. In proposition 4, I show that partners will indeed break-up immediately if and only if there is no friction and that the break-up time is increasing with the friction. I then show that the value of dating in an arranged marriage setting and show that none of this benchmarks is optimal. With these results in hand, we can show the main result of the chapter in proposition 5: the welfare of the partners is not monotonic in the amount of friction.
The intuition behind this result is that increasing the friction has a direct negative effect on the value of dating because if a partner wants to break-up it worsens his situation. But there is also a positive effect, with higher friction partners date longer and increases the value of being in love because it is less likely that the other partner will mistakingly end a good relationship.

In the final section, I discuss some robustness issues of this model. The assumption that partners have common value simplifies the analysis but also makes it less robust to small changes in the model and we discuss this changes in this section.

The game does have asymmetric equilibria. In the first subsection, I discuss an example in which partners are perfectly coordinated between one type that never breaks-up and another that may eventually end the relationship. This case does have a well-defined pure strategy equilibrium and the externality described in the symmetric equilibrium disappears, it is optimal to have no friction whatsoever. Good relationships still might end, but the coordination leads to fewer break-ups so that increasing the friction no longer improves welfare. This result shows that our model is similar to a “chicken game” with one problematic mixed strategy equilibrium and two asymmetric efficient equilibrium.

In the next subsection, we informally discuss some extensions of the model. Since the source of the externality that drives our main result is that players are incapable of communicating their impressions, in the first part of this section, we discuss which extensions would lead to credible communication and in which cases would “cheap talk” prevails. In the second part of that section, we discuss to which extent can we extrapolate the conclusions of our model to other economics institutions.

This chapter relates three different kinds of literature. First we use extensively the methods of strategic experimentation in bandits developed initially by Rothschild [1974] and more recently by Keller et al. [2005], Keller and Rady [2010] and by Murto and Välimäki [2011]. In this line, Bonatti and Horner [2009] and Bonatti and Hörner [2014]
show that moral hazard is a force that leads teams to work too little, too late and to aggregate information suboptimally. McAdams [2010] consider how moral hazard affects stochastic repeated partnership and dating is one of his applications. The conclusions of these models often are qualitatively similar to ours, but the force, different.

In our model partners never break-up with a positive probability, this is different than other papers in the literature, such as Klein and Rady [2011] where agents have a cut-off strategy. The goal of our model is to discuss the optimal dating website, recently other papers have characterized optimal design for different settings, such as Che and Hörner and Damiano et al. [2012]. The last one solves for optimal deadlines in negotiations and has mechanics that are rather similar to ours. We also employ tools from the reputation literature in Board and Meyer-ter Vehn [2013] and specially Bar-Isaac [2003].

There is also natural parallel with the vast literature on job searches\(^4\). In this line, Jovanovic [1979b] is one of the first papers to consider jobs as an experience good and Jovanovic [1979a] works on a model of endogenous break-ups. The main difference to that literature is that in our model, every good match has the same value. Therefore, there are more break-ups than optimal because they are not realizing that they are in good one, and not because they are looking for a better match. In the last section, we discuss the comparison between our model and the job search literature.

Lastly, love used not to be a common topic among economists, Becker [1973] helped to switch the literature and nowadays a substantial part of the economic papers discuss the marriage market. Shimer and Smith [2000] derived sufficient condition for assortative matching when a search has frictions however in our chapter, the surplus of the matched does not vary substantially since there are only good and bad matches\(^5\). As dating moved to the online scene, so did the research in economics, Hitsch et al. [2010] show that online

\(^4\)See Rogerson et al. [2004] for a very useful survey of the job search literature.

\(^5\)This model does have a clear parallel with marriage literature. We could think that instead of dating websites, we have dating market and cities, and we would be looking for the optimal city to find a partner. Although some intuition could be derived from this alternative interpretation, in order to have a coherent story, I will stick to the dating website interpretation.
and offline matches are similar and that they exhibit assortative patterns. They do consider online matching to be non-friction, an assumption that we relax and discuss in this chapter. Finally Mortensen [1988] argues that higher search frequencies would improve the life of partners because fewer would be left unmatched, in our model this would only be the direct effect of a smaller search friction and this is precisely the economic intuition that we would like to sophisticate.

1.2 Model

Two partners $a$ and $b$ are matched together. The match has a common value, it is either good with probability $p_0$ or bad for both of them. The match is an experience good, initially, neither of them knows its quality, time might tell. Time is continuous.

After being matched, they start the dating period. During that time, they receive the flow of utility $f$, regardless of the quality of the match, and discount the future at rate $r$. In addition to this flow, a good match can yield an instantaneous payoff of $v$.

This shock of utility works as news about the quality of the match. It completely reveals its quality and good news will come following an exponential distribution with frequency $\lambda$ and if the match is low quality, these shock will never come.

Conditional on being in a good match these shocks are privately observed and independent. Moreover, the partner cannot credibly communicate if he received the shock of utility. This is critical to the model and the source of externality.

At any time, either partner can end the dating period, if he ends the relationship, they both go back to the dating website and receive the utility $W$. The timeline of the game is described in Figure1.1.
1.2.1 How Beliefs Evolve

Good news completely reveals the quality of the match so the beliefs in this scenario are trivially equal to 1. Let \( p(t) \equiv P(h|t) \) the probability that the quality of the match is high given that they have been dating for \( t \) periods with no news.

Even in the case of a good match, at any period of time, that is a probability that each partner is unaware of its quality since shocks come with an exponential distribution, this probability is given by \( u_t = e^{-\lambda t} \).

Additionally, at every period time, there is an instantaneous probability that the other partner will break-up the relationship, \( b(t) \). Therefore the fact that they are still dating should influence the beliefs about the quality of the match. The Bayes-rule for the evolution of the beliefs is given by:
\[ P_{t+dt} = \frac{p_t e^{-(\lambda + u_t b(t)) dt}}{p_t e^{-(\lambda + u_t b(t)) dt} + (1 - p_t) e^{-b(t) dt}} \]  

(1.1)

1.2.2 The Value of the Website

In this section, I will assume that the value of the website is exogenously determined \( W \). We can think of it as simply being the outside option of the partners. Evidently, the most interesting analysis treats this value endogenously as a result of the optimal behavior of the partners. In section 5, I endogenize this value and set it to be equal to the value of dating at time zero discounted by the friction. In the rest of this section, however, I keep that value as a constant \( W \) since the optimal behavior within a dating relationship should not take into account its impact on the value of the website.

I call arranged marriage the situation when partners are not allowed to break-up. Once they are matched, they are forced to stay together forever. We will call this value, \( A \). This constant is very helpful and will appear many times in our analysis, so it is useful to state the following lemma.

**Lemma 1.2.1.** The value of an arranged marriage is \( A = \frac{f + p_0 \lambda V}{r} \).

In order to see this result, one should notice that the expected value of an arranged marriage is the expected value given that it is a bad relationship plus the expected value given that it is a good.

\[
A = (1 - p_0) \int_0^\infty f e^{-rt} dt + p_0 \int_0^\infty (f + \lambda V) e^{-rt} dt = \frac{f + p_0 \lambda V}{r}
\]

If the partners never break up they will get the flow of utility \( f \) forever, and with probability \( p_0 \) they will get with frequency \( \lambda \) positive shocks of utility \( V \). This value will decrease to \( \frac{f + p_0 \lambda V}{r} \) as time passes without positive news and it will jump to \( \frac{f + \lambda V}{r} \) if good news ever arrives. The above lemma describes this value for \( t = 0 \).
Along this section, we will not be in the arranged marriage scenario, but we will assume that \( W \geq \frac{f + \lambda p_0 V}{r} \), so that the dating website is at least as good as an arranged marriage. We will verify this assumption, later in section 5. We also assume that \( V > W \) so that a partner wants to stay in a good relationship and that \( \lambda > r \), news about the quality of the match comes sufficiently fast.\(^6\)

### 1.2.3 Value Functions

A partner is either informed or not about the quality of the match. We will refer the case of a partner who is aware that he is in a good relationship as in love. If he decides to remain in this relationship he will obtain, \( L(t) \).

\[
L(t) = f dt + \lambda V dt + e^{-r dt} \left[ u_t b(t) W dt + (1 - u_t b(t)) L(t + dt) \right]
\]

(1.2)

The first part is the instantaneous payoff of staying in a good relationship. The second part is the continuation value, with probability \( u_t b(t) dt \) the other partner is unaware that this is good match and will break up in that period and both will be sent back to the dating website and get \( W \) and with probability \( 1 - u_t b(t) dt \) this will not happen and the partner in love will remain in the relationship.

We will call dating a partner who does not know if he is in a good relationship. If he decides to keep dating through the next \( dt \) he expects to obtain:

\[
D(t) = f dt + \lambda p_t V dt + \\
+ e^{-r dt} \left[ \lambda p_t L(t + dt) dt + (1 - p_t (1 - u_t)) b(t) W dt \right] \\
+ e^{-r dt} \left[ (1 - dt (\lambda p_t + b(t) (1 - p_t (1 - u_t))) D(t + dt) \right]
\]

(1.3)

\(^6\)A sufficient condition for the partners to want to say in a good relationship would be \( r W < f + \lambda V \) but the stronger condition imposed here helps to show the existence of an equilibrium.
The first line of the above equation is the expected flow of utility. The second and the third line deal with the continuation values. The second looks at the cases where he received a signal and moves to the value function of the informed type and the case the partner breaks up with him in $dt$ the last line considers the case when nothing happens.

### 1.2.4 Decision Problem and Equilibrium Concept

Maximum principle reduces the dating problem to the optimal choice at every time $t$ for informed and uninformed agents:

$$V_u(t) = \max \{W,D(t)\} \text{ and } V_i(t) = \max \{W,L(t)\}$$

The equilibrium concept will be a Bayes-Nash Equilibrium

**Definition 1.2.1.** An equilibrium in this game consists of a $\langle p(t), b(t) \rangle$ for $\theta = l, h$ such that:

1. beliefs $p_t$ satisfy Bayes rule described in (1)
2. $b(t)$ breaking up rule that solves the maximization problem of each agent.

This breaking up rule can be a mixed strategy. In that case, $b(t)$ will be a hazard rate function for breaking up.

We are also going to rule out weakly dominated strategies because they often yield unconvincing equilibria. For example, if a partner decides to break-up at time $t$ regardless of the information he has received, the other partner is powerless, so he might as well just break-up at the same time, therefore, this strategy, though implausible, would constitute an equilibrium.
1.3 The Analysis

1.3.1 System of Equations

We can rewrite the 3 equations of the model as a system of differential equations.

\[ dp_t = \left( (1 - u_t)b(t) - \lambda \right) p_t (1 - p_t) dt \]  \hspace{1cm} (1.4)

The intuition for this law of motion is that the first term accounts for the fact that a partner should get more optimistic due to the fact that the other did not break up with him, \( b(t) \), weighed by the informativeness of the breaking up, \( (1 - u_t) \), but he should also become more pessimistic because he did not receive a positive shock, \( \lambda \), the later term \( p_t (1 - p_t) \) is the usual dampening factor of beliefs.\(^7\)

The value functions can be simplified by using Taylor Expansion and ignoring terms of order smaller than \( o(dt) \).

\[ L(t) \left( r + u_t b(t) \right) = f + \lambda v + u_t b(t)W + L'(t) \]  \hspace{1cm} (1.5)

\[ D(t) \left( r + p_t \lambda + (u_t p_t + (1 - p_t)) b(t) \right) = f + p_t \lambda v + (u_t p_t + (1 - p_t)) b(t)W + p_t \lambda L(t) + D'(t) \]  \hspace{1cm} (1.6)

1.3.2 Results

It is somewhat intuitive that a partner who is aware that he is in a good relationship will never break up. We state this as a lemma here\(^8\).

**Lemma 1.3.1.** Partner in love never breaks up, \( L(t) > W \).

---

\( ^7\)This law of motion of beliefs is obtained by subtracting \( p_t \) from \( p_t - dt \), diving by \( dt \) and taking the limit as \( dt \to 0 \)

\( ^8\)The proof of this and the subsequent lemmas can be found in the appendices
Before we state the next lemma we should clarify what happens if both partners are in love and that is common knowledge. In that case, since they will never break up, the value of being in love will be \( \frac{f + lv}{r} \equiv L \). This value offers a natural boundary for the value of being in love, so in any situation \( L(t) \leq L \).

**Lemma 1.3.2.** There is no symmetric pure strategy equilibrium.

The idea is that if there were a point in time when partners learn for sure that the other would leave unless he were aware that they are in a good relationship, each partner would find profitable to stay \( dt \) longer and learn what the other knows\(^9\).

If equilibrium exists, there is a phase when the value of dating is higher than the value of going back to the dating website and nobody breaks up, the honeymoon phase. The value function of dating in this phase is described by:

\[
D(t)(r + p_t \lambda) = f + p_t \lambda v + p_t \lambda L(t) + D'(t) \tag{1.7}
\]

During this phase, if no signals arrive, the beliefs and the value of dating slowly decrease until it eventually hits the point where dating is as good as the outside option of going back to the dating website. Thereafter they will be in the break-up phase. As argued above, there cannot be a mass of exit in at that point, so partners not in love randomize their exit time so that the other partner, if not in love, would be indifferent between going back to the dating website or staying dating a bit longer. So during that phase we set \( D(t) = W \) and \( D'(t) = 0 \). Rearranging the value of dating equation (6) we obtain the following equation for the value of being in love in the "break up" phase:

\[
L(t) = \frac{Wr - f}{\lambda p_t} + W - v \tag{1.8}
\]

\(^9\)The statement of the lemma can actually be stronger. The same argument implied that there cannot be break up with any positive probability.
This is the Value of being in Love that would make partners not in love indifferent between continuing dating and breaking up. This equation has important intuitive features that will be pertinent in the rest of our analysis. First, the value of the being in love in the mixed phase is increasing in the value of going back to the dating website. If going to back to the website is not so bad, the value of being in love has to higher so that dating partners are indifferent between breaking up and dating more. Also, since in the "honey-moon" phase the beliefs $p_t$ will unambiguously decrease, the longer the "honey-moon" phase, the higher the value of being in love. We can now differentiate it to obtain:

$$L'(t) = \frac{(Wr - f)(1 - p_t)(\lambda V - b(t)(1 - u_t))}{\hat{\lambda}p_t}$$ (1.9)

We can use both the value function of being in love and its derivative to solve for the hazard rate of breaking up, $b(t)$. We are now ready to state our main result, the construction of the equilibrium in mixed strategies in this game, which summarizes this discussion.

**Proposition 1.3.3.** There exists a unique mixed strategy symmetric equilibrium where:

1. Partners in love never break up.
2. Partners not in love do not break up until $t^*$.
3. After $t^*$ partners not in love break up with hazard rate $b(t)$ where:

$$b(t) = \frac{\hat{\lambda}p_t(f + \lambda V - rW + f + rV - rW) + (Wr - f)(\lambda - r)}{(Wr - f)(1 - p_t(1 - u_t)) - \hat{\lambda}p_t u_t V}$$

Where the time $t^*$ is implicitly defined as the moment when the value of dating is equal to the value of going back to the website:

$$D(t^*) = f dt + \hat{\lambda}p_t v dt + e^{-r dt} [\hat{\lambda}p_t L(t + dt) dt] + e^{-r dt} [(1 - \hat{\lambda}p_t dt) W] = W$$ (1.10)
We obtain the construction of the above result by simply replacing $L(t)$ and its derivative in equation (5) and isolating the break-up ratio. The proof of uniqueness is left for the appendix. We use the unique boundary condition discussed in the next session, then show that the break-up rate and beliefs form a divergent path, so there is a unique equilibrium path.

The statement of the equilibrium also gives us some insight into the question of existence. The hazard rate of break up has to be positive. Our assumption that $V > W$ and that $W_r > f$ assures us that the numerator is always positive, it remains to be tested that the denominator is positive. It must also be true that the induced value function of being on love is higher than the value of going to the website. The condition that $W \geq \frac{f + \lambda p_0 V}{r}$ is sufficient for both the above restrictions. Whenever the honeymoon ends, dating partners are indifferent between breaking up or staying in the relationship.

1.3.3 Properties of the Equilibrium Path and Asymptotic Behavior

How do beliefs evolve in the breaking up phase? What about the value of being in love? Is there any testable implication of this model? In order to answer these questions, we will have first to establish the boundary condition of equilibrium path. The following lemma will be useful.

**Lemma 1.3.4.** *Breaking-up phase never ends.*

I will save the details of the proof to the appendix, but some intuition is useful here. The idea if breaking up phase stops at time $\tilde{t}$, the value of dating will be worst than going back to the website and dating partners will break-up. Thus the belief at time $\tilde{t}$ will be $p_{\tilde{t}} = 1$, so dating partners would like to wait until $\tilde{t}$ for this information to unravel.

The above lemma generates the following boundary condition.
\[
\lim_{t \to \infty} (t) = \bar{L} \text{ and } \lim_{t \to \infty} L'(t) = 0
\]

Since there is always a positive exit of dating partners, those in love are becoming more and more optimistic about the probability of being in a situation where both are in love. Since the unconditional probability of breaking up is diminishing, the value of being in love slowly approaches the theoretical maximum of \( \bar{L} \). This condition allows us to derive the properties of the equilibrium path and to proof its uniqueness.

**Proposition 1.3.5.** During the in the break-up phase:

1. **hazard rate of breaking up is decreasing**
2. **beliefs are decreasing**
3. **value function of being in love is increasing**

We leave the details of the proof for the appendix. The monotonicity of the value of being in love is a direct consequence of the beliefs decreasing and equation 8. We show in the appendix that the hazard rate of breaking up is also a consequence of beliefs decreasing, moreover they could not increase too fast, otherwise, they would force the value of being in love to decrease. The fact that beliefs are decreasing come from our assumption that in the limit beliefs are not changing too much.

From a theoretical perspective, it is interesting to characterize the path of beliefs and the value of being in love, however, those features cannot be observed through data. Break-up rates can. So far we have characterized the conditional break-up rate. A researcher can only observe an unconditional break-up rate. From her point of view, relationships are either good with probability \( p_0 \) or bad with probability \( 1 - p_0 \) and if they are good the probability that the partners is uninformed is given by \( u_t \), therefore, the unconditional hazard rate of break-us is given by:
Finally, we can characterize what happens after partners have been dating for long enough.

**Proposition 1.3.6. Asymptotic Behavior.**

1. \( \lim_{t \to \infty} L(t) = \bar{L} \) and \( \lim_{t \to \infty} L'(t) = 0 \)
2. \( \lim_{t \to \infty} b(t) = \lambda \)
3. \( \lim_{t \to \infty} p(t) = \frac{f(Wr-f)}{\lambda(f+AV+rV-rW)} \)

The first item in the above proposition is just the boundary conditions that we have already been using. The hazard rate of break-ups come directly from the fact that if value function of being in love is not changing, the beliefs have to be constant so \( b(t)(1-u_t) = \lambda \), but asymptotically \( u_t \) goes to zero and we have that \( \lim_{t \to \infty} b(t) = \lambda \). Finally the limit of beliefs comes from the fact that \( \lim_{t \to \infty} L(t) = \bar{L} \). Replacing this expression into equation 8 yields this final condition.

### 1.4 Simulations

So far we have described a few properties of the equilibrium path. We can sharpen our intuition by simulating the model and visualizing the value functions, the belief path and the hazard rate of breaking ups. Since we are simulating the model, we have some degrees of freedom in the choice of parameters. We decided to choose them so that the break-up phase starts at \( t^* = 90 \), the popular 3-month rule\(^\text{10}\). Here are the parameters employed in the following simulations: \( \hat{\lambda} = 0.02, r = 0.01, f = 1, v = 150, w = 140, p_0 = 0.35 \)

\(^{10}\)For an informal evidence of the 3-month rule see http://www.cnn.com/2010/LIVING/11/03/tf.three.month.dating.rule/
Figure 1.2: Value Function of Dating.

Value functions of dating

values

Time

D(t)
The value function of dating is strictly decreasing until $t^* = 90$ when by construction it reaches the value of going back to the dating website and then the breaking up hazard rate keep is such that it keeps its value equal to $W$.

The value function of being in love is always higher than the value of dating and it is decreasing in the non-exiting phase because partners in that phase know that they will enjoy the flow of high utility for sure until $t^*$, but then they will enter in the break-up phase. After $t^*$ the value starts to increase because the probability that the other is also in love is also increasing up and the break-up rate is decreasing. Eventually, it reached the theoretical limit $L = 400$ when it is common knowledge that they are in a good relationship.

The belief an uninformed agent has about the probability of being in a good relationship declines fast in the honeymoon phase due to the absence of news. In the breaking up
Figure 1.4: Beliefs in both phases.

Beliefs in both phases

probability

Time
phase, it is still declining, although not as fast as before because now there is the positive news that the other partner has not broken up yet.

We plot the breaking up rate of partners who are uninformed and the unconditional break-up rates. Although theses curves are not substantially different, it is important to see the unconditional break ups because that is what an external person could observe in a data set. After it starts, the breaking up rates are declining over time, so that eventually only partners in love stay in the relationship and we reach our intuitive boundary condition assumption.

1.5 Welfare

So far we have characterized the equilibrium, its existence conditions, we have described the properties of the equilibrium path and its asymptotic behavior. In this session, we turn our attention to the welfare properties of the equilibrium and derive our main results: that the value of the website is increasing with some amount of friction.

In order to make this analysis meaningful, we no longer can treat the value of the
website, $W$, as an exogenous parameter. We will assume that when the partners break-up, they are sent back to the dating website, and they will eventually resume dating and get the utility $D(0)$. However before being rematched, they have to wait for a period of time that discounts their utilities by $(1 - \gamma)$. $\gamma$ is the friction imposed by the website. Most dating websites do not directly force you to wait a period before start dating, but this can be thought as the consequence of the search restrictions. This time, it could be because the dating website purposely forces the partner to wait a period of time, or simply because it limits the number of visible potential matches so that de facto it takes longer to find a new match. Tinder and Match.com have no restrictions in how many profiles a member can browse, so we can think about them as approaching the non-friction equilibrium. Meanwhile, Hinge, Coffee-Meets-Bagel, and E-Harmony restrain the number of profiles available, in our model, they are choosing a positive friction. The following equation summarizes the above discussion:

$$W = (1 - \gamma)D(0)$$

When $\gamma \to 0$, we are in the case of a non-friction matching, where as soon as there is a break-up the partners immediately dating someone else. Obviously, this is not the case of any real world institutions, which face many restrictions, including the time that takes to browse through dating profiles. The following lemma describes the outcome value of the dating in this framework.

**Lemma 1.5.1.** If partners break-up immediately, the value of the website is $\frac{f + p_0AV}{r}$.

If partners break-up immediately, two things can happen in the first and unique instant of dating: they both fall in love and stay together forever or not. However the probability that they both receive the signal is $p_0\lambda^2 dt^2$. Since we are ignoring terms of order smaller than $dt^2$, this can be approached to zero. Therefore, partners will break-up at $t^* = 0$, be rematched and keep looping between relationships forever gaining $f$ in every relationship.
and with probability $p_o\lambda$ they will get $V$. However, they will never be both in love.

It seems very clear that there a close relation between non-friction matching $\gamma = 0$ and immediate break-ups $t^* = 0$. Indeed, the following proposition formalizes our intuition.

**Proposition 1.5.2. Relationship between $\gamma$ and $t^*$**

1. $t^* = 0$ if and only if $\gamma = 0$

2. $t^*$ is weakly increasing in $\gamma$.

The only if part of the statement is straightforward. If $t^* = 0 \Rightarrow D(t^*) = D(0)$ and this implies that $W = W_1 - \gamma \Rightarrow \gamma = 0$. The if part is slightly more complicated and the second item is an application of implicit function theorem to equation 10. We leave both proofs to the appendix.

The above proposition shows us that friction is crucial to find love. It is the burden of returning to the website that incentivizes the partners stay in the relationship and have a honeymoon phase. Without friction there is no experimentation and the higher the friction, the longer partners will date. The only reason why the relationship is not strict increasing is because the for a friction sufficient high, the partners will never break-up.

what might be bewildering is the fact that being in a dating environment without any friction is as good as being in an arranged marriage at time 0. As we have seen before, partners in a non-friction environment will sometimes gain shock of utilities but will never be in a long lasting relationship. Conversely in the arranged marriage scenario, at any time bigger than 0, to be dating is worse than to be in the non-friction scenario, but there is a possibility that partners will be in love, and that is unambiguously better.

Perhaps it is a good time to discuss what features of the model are crucial for this puzzling point. A model where the news is actually bad will not have this feature, arranged marriages would be worse than never ending dating. However in that model, all
the dynamics of the model would be different and endless dating would be optimal since it minimizes the risk of bad news.

It remains for us to discuss what do we mean by sufficient high friction. The following lemma will put some structure in the discussion.

**Lemma 1.5.3.** *If partners are in an arrange marriage then* $\gamma \geq \bar{\gamma} \equiv \frac{p_0\lambda V}{f + p_0\lambda V}$

The idea of this lemma is that when partners never return to the dating pool, it must be better to keep dating the same person, no matter how bad the relationship is than to go the website. In other terms $\frac{f}{r} \geq \frac{(1-\gamma)(f+\lambda p_0 V)}{\bar{r}}$. Isolating the friction $\gamma$ yields the above result.

We now have to detail what is the value of starting a new relationship $D(0)$ for every possible value of gamma between 0 and $\bar{\gamma}$. Remember that every $\gamma$ is associated with a unique $t^*$. Fix this time, when the honeymoon phase ends, the value of starting a new relationship will be:

$$D(0) = p_0 \left( \frac{(f + \lambda V)}{r} (1 - e^{-rt}) + e^{-rt} \left( (1 - u_t)L(t) + u_t(1 - \gamma)D(0) \right) \right)$$

$$+ (1 - p_0) \left( \frac{f}{r} (1 - e^{-rt}) + e^{-rt} (1 - \gamma)D(0) \right)$$

(1.12)

We have dropped the star from the time the honeymoon phase ends for simplicity of notation. The first line represents the payoff in case of a good match, which happens with probability $p_0$. The first part of the brackets is the expected flow of utility during the honeymoon. The second part of the first line brackets are the two possible continuation values at $t^*$, $L(t)$ in the case that he is in love and $(1 - \gamma)D(0)$, for the case the he is unaware that this is a good relationship. The second line of the equation is the payoff of a bad match. The partner receives the flow of utility until $t^*$ when he gets the discounted continuation value of going back to the dating website.
We are not completely done in our task of rewriting the problem as function of \( D(0) \) because the value of being in love at \( t^* \) was also a function of \( W \) in equation (8) and now it can be rewritten as:

\[
L(t) = \frac{(1 - \gamma)D(0)r - f}{\lambda p_t} + (1 - \gamma)D(0) - v
\]  

(1.13)

Now we have enough structure to state the main result of this section.

**Proposition 1.5.4.** *The value of the website changes non-monotonically with the friction.*

We leave the complete proof of the proposition to the appendix, but we outline the argument here. It relies on the previous results and in an analysis of the value of the website in the arranged marriage scenario when partners never break-up. We have argued that for a friction sufficient high, partners will behave as if they were in this benchmark.

Next we show that for a value of the friction sufficiently high, yet smaller than \( \gamma \) the value of the website is higher than the value of the arranged marriage scenario. This observation and the fact that the value of the website in the arranged marriage is higher than in the frictionless benchmark complete the proof.

In this section, we discussed the main result of this chapter. Some friction can improve the value of dating websites. This occurs because partners cannot communicate their signals and this creates an externality in breaking up.

In the next section, we discuss some of the shortcomings of this model and how more complex models could make this analysis more robust. More precisely, I discuss a few relevant points: is the above result also true in asymmetric equilibria; how reasonable is the assumption that partners cannot communicate their impressions; and to what extent are the conclusions of this model applicable to other situations besides dating and love.
1.6 Discussion

The model has a very useful but unnecessary assumption: common values. In the other extreme, with private values, the fact that partner A is in love tells nothing about the value of the relationship for partner B. Thus forcing them to date longer does not improve the aggregation of information. The intermediate scenario, and probably more realistic, would be to have positively correlated values. Murto and Välimäki [2011] develop a model in this setting. They show that when an agent decides to end the experimentation phase others will exit in waves, but not all of them.

In a similar dating model with correlated values, the main forces of this model would still be present, but the analysis would be less clean. However, many of the shortcomings of the present analysis, discussed in this section, arise because the incentives of partners are perfect aligned and this would not be the case in the model of dating with correlated values.

1.6.1 Asymmetric Equilibrium

So far we have focused on the case of symmetric equilibrium. Now a couple of question bear in mind. Are there asymmetric equilibria? If so, are those equilibria substantially different than the symmetric equilibrium? The short answers are yes and yes.

We will look at one extreme asymmetric equilibrium when there is some perfect co-ordination before the dating period starts and one partner is assigned to the role that never break-up and other ends the relationship optimally.

It is easy to see that this constitutes an equilibrium. The passive partners, if in love does not want to break-up and if not in love, would not want to break-up until the exact point when the active partners breaks-up. It is also important to notice that since both partners will face the same break-up date, and the match has a common value, they both
got the same expected payoff. Therefore at \( t = 0 \), we can simply refer to the value of the website without specifying the role of the partner. At \( t > 0 \) those values differ. The value of being in love and dating for the active partner at \( t \) can be written as:

\[
L^a(t) = \bar{L}
\]

\[
D^a(t) (r + p_t \lambda) = f + p_t \lambda V + p_t \lambda \bar{L} + D'(t)
\]

If the active partner sees the signal and falls in love he is sure that this relationship will last forever since he is the only one who would end it. With this two equations, we can solve for the equilibrium of this game

**Proposition 1.6.1.** This game has a pure strategy equilibrium where:

1. Passive partner never breaks-up.
2. Active partner breaks-up at \( t = \frac{1}{\lambda} \ln \left( \frac{p_0 (\lambda + r)(f + \lambda V - rW)}{(1 - p_0) r(Wr - f)} \right) \)

Differently than in the symmetric equilibrium we can explicitly solve for the optimal break-up time, just setting \( D(t^a) = W \) and \( D'(t^a) = 0 \) yields the above result.

How does the path of the asymmetric equilibrium compare with the path of the symmetric? If they have not observed a signal and are just both dating, the shape of their value functions is similar to the symmetric equilibrium. It starts at \( D(0) \) and for the active partner it slowly decreases to \( W \).

The value function of being in love will be different for the active, for the passive partner and for the symmetric equilibrium. The value of being in love for the active partner is trivially equal to \( \bar{L} \). He knows that he is in love, thus, it is a good relationship and the passive partners will never break-up with him. We simulate the value function of the passive partner when he is in love below. In 1.6, I used the same parameters as in the
simulation section, $\lambda = 0.02, r = 0.01, f = 1, v = 150, w = 140, p_0 = 0.35$ which yields $L = 400$.

The idea is that if the value of been in love slowly decreases as the $t \to t^a$ because, at the break-up point, there is a probability that the active partner has not seen the signal and will break-up. If he does not, it must be that he has seen it and they will be together forever, thus the value function jumps to $L$.

If is important to notice that in the above simulation $t^a = 118$ which is greater than 90 days that was set in our symmetric simulation. Is this just a particular case or a deeper property of the asymmetric equilibrium. The following proposition will clarify this issue.

What about the welfare properties? Is the value of the website in this asymmetric equilibrium also non-monotonic in the amount of friction? Can we say that the non-friction website is sub-optimal? The following proposition will hopefully settle this debate.

**Proposition 1.6.2.** In the asymmetric equilibrium:

1. $t^* = 0$ if and only if $\gamma = 0$

2. $t^*$ is weakly increasing in $\gamma$. 

28
3. The non-friction website value is \( L \).

The first two lines of the above proposition are similar to the symmetric equilibrium. The value of the website, however, is remarkably different than the value in the symmetric equilibrium. The details of the above proof can be found in the appendix, but the following argument is intuitive. The value of the website in this equilibrium can be written as:

\[
D^a(0) = p_0 \lambda L dt + (1 - p_0 \lambda dt) (1 - \gamma) D^a(0) \tag{1.16}
\]

Setting \( \gamma = 0 \) and isolating the value of the website, yields the above result.

Can we generalize this result? Can we compare the values of the dating website in the asymmetric equilibrium with the symmetric, for all values of the friction? The answer as we can see in the following proposition is yes.

**Proposition 1.6.3.** For all \( \gamma \), \( D^a(0) \geq D(0) \)

So the value of the website in the asymmetric equilibrium is not only higher for the non-friction website, but it is higher for any level of friction. I leave the details of the proof to the appendix, but I can sketch the argument here. First I show that for any fixed break-up time, the value of the asymmetric equilibrium is higher than the symmetric. Moreover, I show that the break-up time \( t^a \) maximizes the value of the website and this completes the proof.

### 1.6.2 Communication

The main result of the chapter is that the optimal dating website might have to include some friction. From a theoretical perspective, this is not new. The novelty in this chapter is that the source of externality is the fact that players are incapable of communicating their signals. Thus, a reasonable question would be: why do players do not communicate?
A fair answer would be: if they could, they might. Let us clarify that. First, if players were allowed, but this communication was imperfect, noisy, the model would be more complicated because instead of two ordinary differential equations as value functions, we would have a partial differential equation with time and signals as states. However, all the forces that currently play a role in the model would still be present and we can imagine that the equilibrium and welfare property would have similar qualitative features.

Second, if players are allowed to communicate, to truthfully report their impressions would be an equilibrium. Their interests are completely aligned and by doing so, they completely aggregate the information available. However ”cheap-talk” would also be an equilibrium. To see that notice that the following line of reasoning should apply: given that my partner is telling me the truth about his signals and I know my signal, I am indifferent between telling him the truth or lying”. In particular, the partner could say that he has observed the signal and he is in love, when he is not. Moreover, if we introduce any arbitrary small misalignment of interests, the ”cheap-talk” equilibrium would be the only one to remain. There are many features of relationships that justify contradictory interests. For example, if having your partner breaking up with you is worse than breaking up with him, all partners would have incentives to lie, to say that they are in love when they are not, just to be the one doing the crucial decision. Similarly, if the flow of utility of being in a relationship is higher when the other thinks that you love him, lying will prevail.

Therefore, although communication seems to be a natural extension of the model, noisy communications should not change the qualitatively aspects of the equilibrium and the idea the truth would emerge naturally is not clear and not robust to small asymmetrical incentives.

\[^{11}\]For a comprehensive survey of the ”cheap talk” literature see Crawford [1998]
1.6.3 Relationship with other economic institutions

Hopefully, the sentences in the beginning of the introduction were sufficient to convince the reader that dating websites are an important feature of modern life and should be studied in themselves. However we acknowledge that they are not the bread-and-butter of economists, so we want to discuss to which extent our conclusions apply to other economic institutions.

Our model is essentially about aggregating information when communication is not possible or allowed. We employed the assumption of common values for simplicity, but the properties of the model should hold when values are correlated as in Murto and Välimäki [2011].

Job market search is a natural parallel to dating websites. The players would be the firms and the employees and dating would be the internship or the interview process before they commit to full-time employment. The direct extrapolation of our conclusions is that an agency that helps job seekers should offer a limited number of target job options so that employees and firms could truly experiment in the job and discover whether they are a good fit.

However, we have to proceed cautiously. Money is completely absent in our model and in the previous section discussion. Even if communication is not possible or credible, players can signalize their impressions with cash-transfer. This feature is often present in some instances of the job market when hiring bonus or moving expenses compensations are available, but it is absent in others: graduate students probably would accept to teach for free in a top university, just to show how good fit they are, but we are not allowed to.

Whether dating partner actually can credibly signal his feelings with money is an interesting topic of debate that would mitigate our conclusions. Summarizing, our model extends naturally to economics situations where there is common value, an experience good, and cash transfers are not allowed.
1.7 Mathematical Appendix

Proof of Lemma 2. Since the partner always have the options to go back to the dating website we can write his value function as:

\[ L(t) = (f + \lambda v)dt + e^{-r dt} [u_t b(t) W dt + (1 - u_t b(t)dt) \max \{L(t + dt), W\}] \]

\[ \geq (f + \lambda v)dt + e^{-r dt} [u_t b(t) W dt + (1 - u_t b(t)de)W] \]

\[ = (f + \lambda v)dt + e^{-r dt} W \approx (f + \lambda v)dt + (1 - rdt)W = (f + \lambda v - rW)dt + W > W \]

□

Proof of Lemma 3. Assume there is \( t^* \) such that both partners break up for sure if they are not in love. We have shown that they will never break up if in love, so the decision to keep dating for \( dt \) longer at \( t^* \) yields:

\[ D(t^*) = (f + p_t \lambda v)dt + e^{-r dt} [p_t u_t L(t + dt) + (1 - p_t u_t)W] \]

\[ = (f + p_t \lambda v)dt + e^{-r dt} \left[ p_t u_t \frac{f + \lambda v}{r} + (1 - p_t u_t)W \right] \]

Taking the limit as \( dt \to 0 \)

\[ D(t^*) = p_t u_t \frac{f + \lambda v}{r} + (1 - p_t u_t)W > p_t u_t W + (1 - p_t u_t)W = W \]

□

Proof of Lemma 4. Let’s first establish that if the breaking-up phase ends, it never resumes. To see this notice that during this phase:

\[ dp_t = (b(t)(1 - u_t) - \lambda) \]
So if \( b(t) = 0 \) the beliefs will decrease. Furthermore, notice that in the following equation, the break-up rate is strictly increasing in the belief.

\[
b(t) = \frac{\lambda p_t (f + \lambda V - r W + f + r V - r W) + (W r - f)(\lambda - r)}{(W r - f)(1 - p_t (1 - u_t)) - \lambda p_t u_t V}
\]

So, since \( p_{t+\varepsilon} < p_t \) we will have that \( b(t + \varepsilon) < 0 \) which is not possible. Thus we have established that if dating partners stop breaking up, they will never resume. Now noticed that during the breaking-up phase the value of dating can be written as:

\[
D(t) = \frac{f + \lambda p_t V + (u_t p_t + (1 - p_t))b(t) W + p_t \lambda L(t)}{r + p_t \lambda u_t b(t) + (1 - p_t)b(t)}
\]

But if \( b(t) \) is 0 forever and \( \lim_{p_t \to \infty} = 0 \) the above expression becomes:

\[
D(t) = \frac{f}{r} < W
\]

Which contradicts the indifference condition \( \square \).

\[\Box\]

Proof of Proposition 1. The construction of the equilibrium was argued before, it remains for us to show that this equilibrium the exists and that it is unique.

Let’s first discuss the existence condition. There are two restrictions on the existence. The break-up rate has to positive and it cannot be so high that partners in love would rather go back to the dating website. We will talk about each of this conditions.

The break-up rate is given by:

\[
b(t) = \frac{\lambda p_t (f + \lambda V - r W + f + r V - r W) + (W r - f)(\lambda - r)}{(W r - f)(1 - p_t (1 - u_t)) - \lambda p_t u_t V}
\]

The numerator will always be positive, we have to be sure that the denominator is also positive.

\[
(W r - f)(1 - p_t (1 - u_t)) \geq \lambda p_t u_t V \Leftrightarrow \frac{W r - f}{(1 - u_t)(W r - f) + u_t \lambda V} \geq p_t = \frac{p_0 u_t}{p_0 u_t + 1 - f}
\]

33
The last equality comes because before the breaking up phase beliefs only change due to the absence of signal. Rearranging the terms of the inequality we obtain:

\[
u_t^2 \leq \frac{(W_f - f)(1 - p_0)}{(f + \lambda v - W_r)p_0}
\]

Since \(u_t = e^{-\lambda t}\) the above equation can be rewritten in terms of \(t^*\):

\[
t^* \geq \frac{1}{2\lambda} \ln \left( \frac{(f + \lambda V - W_r)p_0}{(W_r - f)(1 - p_0)} \right)
\]

The other condition that must be satisfied in that \(L(t^*) \geq W\). That is:

\[
L(t) = \frac{W_f - f}{\lambda p_t} + W - v \geq W \iff \frac{W_f - f}{\lambda v} \geq \frac{p_0 u_t}{p_0 u_t + 1 - P_0}
\]

Isolating \(u_t\) in the last inequality we obtain:

\[
u_t \leq \frac{(W_f - f)(1 - p_0)}{(f + \lambda V - W_r)p_0} \iff t^* \geq \frac{1}{\lambda} \ln \left( \frac{(f + \lambda V - W_r)p_0}{(W_r - f)(1 - p_0)} \right)
\]

Clearly, the first condition we obtained implies the second. We have only solved for \(t^*\) implicitly. However \(t^*\) has to be positive, thus for any \(t^*\), these conditions will not be binding if the right side of the above inequality is negative and this will occur precisely if \(W \geq \frac{f + p_0 \lambda V}{r}\), which is our condition that the value of the website is higher than the value of the arranged marriage \(\square\).

We shall not turn our attention to the uniqueness issue.

We have a system of differential equations formed by \(b(t)\) and \(p_t\). Assume for some \(t > t^*\) there are two possible beliefs in different equilibrium paths \(p_{1t}\) and \(p_{2t}\), WLOG \(p_{2t} > p_{1t}\). For the same \(t\), \(b(.)\) is strictly increasing in \(p_t\), thus \(b_2(t) > b_1(t)\). In addition, notice that \(dp_t = (b(t)(1 - u_t) - \lambda)\) is also increasing in \(b(t)\). Therefore \(dp_{2t} > dp_{1t}\). So we have:
\[ p_{2t+\varepsilon} - p_{1t+\varepsilon} \approx p_{2t} + dp_{2t} - p_{1t} - dp_{1t} > p_{2t} - p_{1t} \]

Since the \( t \), \( p_{1t} \) and \( p_{2t} \) were arbitrarily chose, those are divergent paths and they cannot have the same boundary condition. \( \square \)

**Proof of Proposition 2.** We first show that if \( dp_t < 0 \) then \( b'(t) < 0 \). Differentiating \( b(t) \) we obtain:

\[
b'(t) = \frac{\lambda dp_t (2f + l\lambda V + \lambda r - 2rW) \text{den}}{\text{num}} + dp_t \left(1 - u_t\right)(Wr - f) + u_t \lambda f \]

\[\text{den} + dp_t \left(1 - u_t\right)(Wr - f) + u_t \lambda f \text{num}
\]

\[\text{num} - p_t \lambda u_t (f + \lambda V - Wr).\text{num}
\]

(1.17)

Where num and den are the numerator and the denominator of the breaking up rate. The numerator is always positive and the denominator has to be positive whenever an equilibrium exists. Clearly, if \( dp_t < 0 \) then \( b(t) \) will be decreasing.

Now assume that for some interval \( t \) to \( t + \varepsilon \) in the breaking up phase \( b(t) \) is increasing, the converse of the above conclusion, this implies that \( p_t \) is increasing. But this only happens if and only if \( b(t)(1 - u_t) > \lambda \). Since \( b(t) \) is an increasing function and so it is \( (1 - u_t) \) this condition will forever hold and the limit of \( p_t \) and the value of \( L(t) \) will be different than our stipulated boundary condition. Therefore, \( b(t) \) has to be decreasing.

But since \( u(t) \) is always decreasing and so is \( b(t) \) in the breaking up phase, from equation 5 and our boundary condition that \( \lim_{t \to t} L(t) = \bar{L} > L(t) \) the value function of being in love \( L(t) \) has to be increasing. However, equation 8 shows us that this will only occur if \( p(t) \) is also always decreasing. \( \square \)

**Proof of Proposition 4.** We argued the if statement in the body of the text. It remains for us to argue the other direction and the monotonicity condition.
If $\gamma = 0$ then $D(0) = W \equiv D(t^*)$, but that implies that the value of the dating website is equal to the value of breaking up immediately, $\frac{f}{r}$. However we have to establish that $t^* = 0$. We will do this by showing that the derivative at 0 is negative. Thus if the partner is indifferent between breaking up and continuing dating at $t = 0$, he will strictly prefer to break up at $t = dt$. To see this we need to check that equation (5) and (6) in the honeymoon phase become:

$$D(t)(r + \lambda p_t) = f + p_t \lambda V + p_t \lambda L(t) + D'(t)$$

Rearranging this equation and isolating $D'(t)$, we obtain.

$$D'(t) = D(t)(r - f - \lambda p_t V - \lambda p_t (L(t) - D(t)))$$

When $t = 0$ and $D(0) = \frac{f + p_0 \lambda V}{r}$, this equation becomes:

$$D'(0) = -\lambda p_0 (L(0) - D(0)) < 0$$

Where the last inequality comes from the fact that $L(t) > W$. So if partners are indifferent between breaking up or keep dating at $t = 0$, they will break-up immediately $dt$ later, therefore, $t^* = 0$.

We know what to show the monotonicity condition. Notice that equation (10), that implicitly defines $t^*$, can be written as:

$$\Delta = f dt + \lambda p_t V dt + e^{-\lambda p_t dt}(\lambda p_t L(t^*) dt + (1 - \lambda p_t dt)W) - W = 0$$

Using the fact that $L(t^*) = \frac{WR - f}{\lambda p_t} + W - V$, we can rewrite the above equation as:

$$\Delta = f dt + \lambda p_t V dt + e^{-\lambda p_t dt}((1 + \lambda p_t dt)W - f dt - \lambda p_t V dt) - W = 0$$
Implicit function theorem tell us that:

\[ \frac{\partial t^*}{\partial \gamma} = -\frac{\partial \Delta}{\partial t^*} \]

Let's first work on the derivative in the denominator:

\[
\frac{\partial \Delta}{\partial t^*} = (f + \lambda p_t V) dt^2 - \lambda^2 p_t (1 - p_t) V dt - re^{-rdt} ((1 + r dt) W - f dt - \lambda p_t V dt) \\
- e^{-rdt} ((f + \lambda p_t V) dt^2 - W r dt^2 - \lambda^2 p_t (1 - p_t) V dt) \\
\approx -\lambda^2 r p_t (1 - p_t) V dt^2
\]

(1.18)

The first equality is just the derivative of the above equation, the final approximation comes from ignoring terms of order higher than \(dt^2\). Now let us work with the numerator, but first it is worth remembering that \(W = (1 - \gamma) D(0)\).

\[
\frac{\partial \Delta}{\partial \gamma} = \frac{\partial \Delta}{\partial W} \cdot \frac{\partial W}{\partial \gamma} = D(0) r dt^2
\]

The above equation is positive since \(D(0) \geq \frac{f}{\tau} > 0\). Therefore we have:

\[
\frac{\partial t^*}{\partial \gamma} = \frac{D(0)}{\lambda^2 p_t (1 - p_t) V} > 0 \square
\]

\(\square\)

**Proof of Proposition 5.** We have shown previously that the value of an arranged marriage is equal to the frictionless website. In other words, the value of the website when \(\gamma = 0\) is equal to the value of the website when \(\gamma \geq \gamma > 0\). The remaining of the proof consists in showing that there exists a friction sufficient high but smaller than \(\gamma\) so that the value of the website is larger than the value of arranged marriage. This is sufficient for non-monotonicity.
Following Lemma 1 the value of arranged marriage is:

$$\lim_{\gamma \to \bar{\gamma}} D(0) = \frac{f + p_0 \lambda v}{r}$$

We now have to argue that exists a value of friction so that partners are better than in the arranged marriage setting. Similar we have the condition that $L(t) \leq \bar{L}$. Thus, since $0 < D(0) \leq \bar{L}$ we can define the limit belief for which it is feasible to have break-ups, $p$ as:

$$\frac{(1 - \gamma)D(0)r}{\lambda p} + (1 - \gamma)D(0) - \nu = \bar{L}$$

Now let’s define:

$$\tilde{\gamma} \equiv \frac{\lambda v(p_0 - p)}{f + \lambda p_0 v}$$

Since beliefs are decreasing in the honeymoon phase, the above friction is well defined, $0 < \tilde{\gamma} < \bar{\gamma}$. So for that friction the agent will have well-defined honeymoon and break-up phases.

Since partners in love will never break-up, the following inequality always holds: $L(t) > (1 - \gamma)D(0)$. In particular this is true for $t^*$, so we have:

$$\frac{(1 - \gamma)D(0)r - f}{\lambda p_t} + (1 - \gamma)D(0) - \nu > (1 - \gamma)D(0) \iff D(0) > \frac{f + \lambda p_t v}{r(1 - \gamma)}$$

This will hold for the value of dating at $t = 0$ regardless of $\gamma$. What is the relationship between the above expression and the value of arranged marriage?

$$\frac{f + \lambda p_t v}{r(1 - \gamma)} \geq \frac{f + \lambda p_0 v}{r} \iff \gamma \geq \frac{\lambda v(p_0 - p_t)}{f + \lambda p_t v}$$

Since the beliefs are decreasing in the honeymoon phase, this inequality will hold for any $\gamma \in (\tilde{\gamma}, \bar{\gamma})$. So in this case, in particularly for "break-up" phases that start sufficiently
late, the value of the website will be higher than the value of arranged marriage, together
with the proposition that the value of dating in the frictionless model is increasing with
the friction, we can conclude that the value of website to be non-monotonic in the friction □.

Proof of Proposition 7. The first item of the proof is very similar to the proof of the sym-
metric equilibrium. If \( t^* = 0 \) then \( D(0)^a = (1 - \gamma)D(0)^a \Rightarrow \gamma = 0 \). Also, the derivative of
the value of dating is negative so \( D(dt)^a < D(0)^a \) and partners will want to break up at
\( t^* = 0 \).

The second item of the proof can be shown by simply differentiating \( t^* \) with respect
to \( \gamma \).

Finally, as described in the body of the text, in the non-friction website the partnership
will end at \( t^* = 0 \). So its value can be written as:

\[
D^a(0) = p_0\lambda Ldt + (1 - p_0\lambda dt)(1 - \gamma)D^a(0)
\]

(1.19)

Setting \( \gamma = 0 \) and isolating the value of the website, yields the last item’s result □.

Proof of Proposition 8. For now on the superscript \( a \) will always refer to the asymmetric
equilibrium. In order to clarify the argument we should write the value of dating as a
function of time and its break-up time, \( D(0,t^*) \) Let us first show that the for any fixed
break-up date \( t^a = t^* \) the value of the asymmetric equilibrium is higher, that is \( D(0,t) \leq
D^a(0,t) \). We know from equations 7 and 15 that:

\[
D(t,t^*)(r + p_t\lambda) = f + p_t\lambda v + p_t\lambda L(t) + D'(t,t^*)
\]

\[
D^a(t,t^a)(r + p_t\lambda) = f + p_t\lambda V + p_t\lambda L + D'^a(t,t^a)
\]

Setting the above equations at the break-up point we have that \( D(t^*,t^*) = W, D^a(t^*,t^a) = W^a \) and \( D'(t^*,t^*) = D'^a(t^a,t^a) = 0 \) and rearranging we have got:
\begin{equation*}
W = \frac{f + \lambda p_t (V + L(t))}{r + \lambda p_t}
\quad \text{and} \quad
W^a = \frac{f + \lambda p_t (V + \bar{L})}{r + \lambda p_t}
\end{equation*}

Since we know that \(L(t) \leq \bar{L}\), for any fixed break-up date \(W \leq W^a\) and therefore \(D(0) \leq D^a(0)\). We have to show that this is also true for different break-up dates. We shall work now with the explicit value of \(D^a(0, t)\), this can be written as:

\begin{equation*}
D^a(0, t) = \frac{f + p_0 \lambda V}{r} \left(1 - e^{-rt}\right) + e^{-rt} p_0 u_t \bar{L} + e^{-rt} (1 - p_0 u_t) W^a
\end{equation*}

The first part of the above equation is the expected flow of utility until the break-up time \(t\) the second and third are the continuation values in the case he falls in love and in the case that they go back to the website. Differentiating the above function we can solve for the optimal break-up time:

\begin{equation*}
t_{\text{max}} = \frac{1}{\lambda} \ln\left(\frac{p_0 (\lambda + r) (f + \lambda V - rW)}{(1 - p_0) r (Wr - f)}\right)
\end{equation*}

Which is the same break-up time the agent chooses, \(t_{\text{max}} = t^a\). We can indeed verify that at this point, the second derivative is negative and it is indeed a maximum. Since this is the unique solution for the maximization problem. We have that at the time \(t^*\) that characterizes the break-up time for the symmetric equilibrium \(D(0, t^*) \leq D^a(0, t^*) \leq D^a(0, t_{\text{max}}) = D^a(0, t^a)\) and that concludes our proof \(\square\).
CHAPTER 2

Identification of Beliefs on Decision Making

2.1 Introduction

If we observe an individual making decisions and if we also observe the signals that he has received, what can we say about his beliefs? In this chapter, we direct the tools of identification to address this question. Whereas the classical approach of decision theory assumes a rich set of choice problems, we work on a binary choice problem, assuming a potentially rich set of signals to achieve identification.

Recent literature has found heterogeneous beliefs to be associated with the formation of bubbles, Xiong [2013]; motivation to commit crimes, Lochner [2003]; and the choice of contraceptive methods, Delavande [2008]; among others, see Manski [2004]. In many decision problems, the agent’s beliefs cannot be disentangled from his utility. However in the above models, beliefs, not utilities, are the most interesting object, as they help to explain past actions and predict future behavior.

In our model, we address situations where beliefs are unknown, yet restrictions can be placed on agent’s utility. Our work shows that in this setting, if the space of signals is sufficiently rich, an econometrician can precisely identify any belief the decision maker may have. We initially achieve this result when there are only two states in the world. Subsequently, we define richness of signals in multiple finite states, show that this condition is necessary for identification of beliefs, and we briefly discuss the sufficient conditions. Following the literature on partial identification developed in Manski [2003], we also de-
scribe what an econometrician can say when there is not sufficient information to point identify beliefs. In these cases, such as when the information has limited likelihood or when there is a finite number of signals, we construct the sharpest bounds an econometrician can form around the agent’s beliefs. We also analyze the case when the agent is dynamically updating his beliefs in a sequence of decisions and show how to construct the sharpest boundaries around his priors.

We first direct our attention to equilibrium beliefs. In this setting, a single agent makes a binary decision after observing a signal which allows him to update his belief about his current state. The econometrician observes the signals and the decisions made repeatedly\(^1\). Based on this information, she makes an inference about the agent’s belief. We show that if the state is relevant, the space of signals is rich enough, and the beliefs are non trivial, we can identify any equilibrium belief the agent might have in a model with two states. We generalize this result for the case of a finite number of states and discuss how rich the distribution of signals must be to achieve identification. In order to derive the above results we do not impose any meaningful assumptions on the agent’s utility function. If we consider that the econometrician also knows the decision rule, which is a function of the utilities in each state, we show that even if the space of signals is not rich, there will be a region where the beliefs are point identified and another region where they can be partially identified. Moreover, we show how one can impose non-trivial bounds on the agent’s beliefs even when there is a finite number of possible signals.

The equilibrium beliefs approach, although closely related to the identification literature, may be unsatisfactory in many economic situations. For this reason, our second approach addresses identification of priors that are updated after each signal. For a given

\(^1\)We focus our work in the individual interpretation of the decision making process, when we observe an individual making the same decision repeatedly. However, the results of this chapter, specifically those in the next section, can be understood in the context of a population of identical individuals receiving conditionally independent signals and the variable of interest would be the average belief in the population. Although this could be a more credible data set, we refrain from getting back to this interpretation for the sake of having a coherent story.
sequence of signals, if we know the conditional distribution of signals, non-degenerate priors hold a one-to-one relation with posteriors. Using this relationship, we show that with any finite number of observations, we can impose non-trivial bounds on the agent’s priors. The bounds are constructed by looking at the decisions made when the agent is least sure about the optimal decision for their current state. Since we are looking at these marginal decisions, the sharpest bounds will depend on the entire history of signals and decisions. As such, to best infer the agent’s priors and current beliefs, the econometrician should be concerned with all past observations.

This chapter follows the tradition founded by Ramsey [1931] and Savage [1972], who describe probability as a subjective feature based on the actions they could generate. According to Savage’s view, if we could offer the agent a menu of lotteries, we could find the one which makes him indifferent between actions. From this it would be possible to recover the probability he assigned to each state of the world. We only offer the agent binary choices but infer these probabilities from how the choice changes as the agent receives information. In addition to these seminal works, Anscombe and Aumann [1963] provided a proof of the existence of these subjective probabilities and Gul [1992] showed that we can guarantee their existence in a setting with a finite number of states. More recently, Ross [2013] showed that observing prices from all possible states, we can recover both utilities and beliefs. Arieli and Mueller-Frank [2014] show that if the action set is uncountable, then there exists a continuous utility function such that actions reveal beliefs.

Our work is also closely related to the literature on experts theory and scoring rules. Savage [1971] formalized the discussion on how to elicit probabilities. More recently, Dillenberger and Sadowski [2012] and Lu [2013] discussed identification on decision theory models when we can observe the the individuals’ entire preference relation or the entire stochastic choice function respectively. Dillenberger and Sadowski [2012] describes conditions on utilities over menus of securities which correspond to an individual behaving
as if they had a distribution over the probabilities of the outcomes and was choosing optimally. In Lu’s set up, the econometrician observes a probability distribution over the menu of choices and he determines whether these choices are consistent with the realization of a random private signal. He shows that choices among binary decisions are sufficient to recover information received by the agent if we employ a test function, whose known payoff varies. In our model, the choice space is restricted to the simplest example of a binary decision where the payoff does not vary, so the identification question depends on the richness of the signal space.

An alternative approach to inferring beliefs is to directly elicit them from agents. Nyarko and Schotter [2002] show how to elicit beliefs in a game using proper scoring rules; Delavande [2008] survey sexually active women in Chicago and elicited expectations of alternative contraceptive methods; Lochner [2003] employ data about expectations of being arrested to construct a model of utility for criminal behavior; and Delavande et al. [2011] review the methods of eliciting beliefs in developing countries and argue that probabilistic questions can be accurately answered even by individuals with very low education levels. Manski [2004] offers a comprehensive survey of this literature, defending the survey based approach for determining beliefs. This approach can suffer from misreporting due to inability or unwillingness to give correct answers. We believe that results should be complemented by and consistent with inference of beliefs from choice data.

The rest of this chapter is organized as follows. Section 2 describes the general model and derive the results for the identification of beliefs. Section 3 uses a modified model to address the issue of identification of priors. Section 4 concludes.
2.2 Identification of Beliefs

The decision making problem  There is one agent making a binary decision in an uncertain world with two possible states.\(^2\) The agent does not know which world he is in, but has a belief about the likelihood of each state. Let the decision of the agent be whether or not to invest \((d = i \text{ or } ni)\) in a world that is either good or bad for investment \((s = h \text{ or } \ell)\).\(^3\) Before making his decision, the agent receives a signal \(x \in X\), say available economic indicators, which allows him to update his beliefs from known conditional distributions, \(f(x|h)\) and \(f(x|\ell)\).

We will incorporate the uncertainty in our model in two different ways. In the first one, which we consider throughout this section, there is an independent identically distributed move of nature in each period. Therefore a new state is drawn before each signal is received and decision is made. In this case the belief of the agent, denoted by \(p = (p(h), 1 - p(h))\), can be interpreted as the agent’s belief about the distribution of this move of nature, and it remains fixed for each decision. Therefore the agent only considers the signal realized in the current decision process when deciding which action to take. In the next section we will allow the agents beliefs to be evolve as signals are aggregated.

We restrict the signal space to be countable but potentially infinite in every state of the world. Although informative about the state, the signal does not directly influence the agent’s payoff. After observing it, the agent updates his beliefs according to Bayes’ rule and his posterior probability is given by:

\[
p(h|x) = \frac{f(x|h)p(h)}{f(x)} = \frac{f(x|h)p(h)}{f(x|h)p(h) + f(x|\ell)(1 - p(h))}
\]

(2.1)

His payoffs are characterized by the values, \(v(i,h), v(ni,h), v(i,\ell),\) and \(v(ni,\ell)\), the payoff of each decision in each state. We define \(\tilde{v}(s) \equiv v(i,s) - v(ni,s)\), the difference between

\(^2\)In a later subsection we will discuss our results for a finite number of states \(S\).
\(^3\)In this sense a binary decision is the hardest case since working with more decisions would make the identification strategies easier since every decision would provide more information.
investing and not investing in each state. In order to keep the decision problem from becoming trivial, we make an assumption on these payoffs.

**Assumption 2.2.1.** \( \tilde{v}(h) > 0 > \tilde{v}(\ell) \)

This condition implies that the agent wants to invest in the high state and does not want to invest in the low state. Since he does not know in which state he is, he will invest exactly when

\[
E_{p(s|x)}[\tilde{v}(s)] \geq 0.
\]  

(2.2)

**The econometrician problem** The econometrician sees the same signals as the agent and she observes his decisions. She also knows the conditional distribution of signals, \( f(x|s) \), so every observation can be summarized as a pair \((d_i, x_i)\). Her main objective is to identify the agent’s initial beliefs \( p \). We follow the notation and definition of identification in Manski [2003].

**Definition 2.2.1.** The set of priors that is observationally equivalent to \( p \) is given by

\[
H[p] \equiv \left\{ p' \in \Delta \mathbb{R}^S : \{d(p, x_i)\}_{i=1}^{\infty} = \{d(p', x_i)\}_{i=1}^{\infty} \right\}.
\]

Each \( p' \in H[p] \) is consistent with the decisions of an agent with prior \( p \) for any information that could be received. If \( H[p] \) is a strict subset of \( \Delta \mathbb{R}^S \), \( p \) is partially identified, and if \( H[p] \) is a singleton, then we say that \( p \) is point identified. If \( H[p] \) is a singleton for each \( p \), then we say the beliefs of the agent are point identified.

**Identification result** We can manipulate equation (2) to characterize when the agent will invest given his posteriors and payoffs.

\[
p(h|x) \geq \frac{\tilde{v}(h)}{\tilde{v}(h) - \tilde{v}(\ell)} \equiv \bar{d}
\]  

(2.3)
Hereafter $\bar{d}$ can be interpreted as the threshold posterior of the decision rule, where the agent is indifferent between investing and not investing. Combining equations (1) and (3), we can write the decision rule in terms of the agents beliefs.

$$\frac{f(x|h)p(h)}{f(x|h)p(h) + f(x|\ell)(1 - p(h))} \geq \bar{d} \quad (2.4)$$

We will use the likelihood ratio as a summary statistic for the information that is observed by the agent and the econometrician. This likelihood ratio is defined by

$$\gamma(x) \equiv \frac{f(x|h)}{f(x|\ell)} \text{ for each } x \in X.$$ 

We can now write the decision rule for the agent in terms of his prior belief about the distribution of states, our variable of interest. The agent chooses to invest given signal $x$ exactly when

$$p(h) \geq \frac{\bar{d}}{\bar{d} + \gamma(x)(1 - \bar{d})}. \quad (2.5)$$

In order for the econometrician to point identify the agent’s beliefs there must be enough information available. Not only must there be information that can sway both an optimistic and pessimistic agent, this information must also be dense enough to differentiate beliefs that are relatively similar. These requirements are formalized in the following full range assumption on the likelihood ratio of the signals in the information set.

**Assumption 2.2.2.** For each $y \in (0, \infty)$ and $\varepsilon > 0$ there exists a signal $x$ such that $\gamma(x) \in B_\varepsilon(y)$.\(^5\)

An information set which satisfies the full range assumption is sufficient for the econometrician to identify the agent’s beliefs. This is shown in the following result.

**Proposition 2.2.1.** If the state is relevant and likelihood ratio of the signals has full range then the initial beliefs of the decision maker are point identified.

\(^4\)Even if the econometrician does not know agent’s utilities, assumption 1.1 is sufficient to guarantee that $\bar{d} \in (0, 1)$.

\(^5\)Since the rationals are countable and dense on the real line, any information set $X$ which $\gamma(\cdot)$ maps onto the rational numbers would satisfy this property.
Proof. Let \( p = (p(h), 1 - p(h)) \) and \( p' = (p'(h), 1 - p'(h)) \) be two distinct beliefs. Without loss of generality assume that \( p(h) > p'(h) \). We define \( \hat{p} \) and \( \hat{y} \) such that

\[
\hat{p} = \frac{p(h) + p'(h)}{2} \quad \text{and} \quad \hat{y} = \frac{(1 - \hat{p})d}{\hat{p}(1 - d)}
\]

Since \( \hat{p} \in (0, 1) \) and \( \overline{d} \in (0, 1) \) we have that \( \hat{y} \in (0, \infty) \). By the full range assumption, there exists an \( \hat{x}(\varepsilon) \) such that \( \gamma(\hat{x}(\varepsilon)) \in B_{\varepsilon}(\hat{y}) \), for all \( \varepsilon > 0 \). For small enough \( \varepsilon > 0 \), we have that

\[
p(h) > \frac{\overline{d}}{\overline{d} + \gamma(\hat{x}(\varepsilon))(1 - \overline{d})} \approx \hat{p} > p'(h).
\]

With the signal of \( \hat{x}(\varepsilon) \), \( d(p, \hat{x}(\varepsilon)) = i \) and \( d(p', \hat{x}(\varepsilon)) = ni \). This shows that the two beliefs \( p \) and \( p' \) are not observational equivalent. Since \( p \) and \( p' \) are arbitrary beliefs of the agent, all beliefs are identified.

This result closely resembles the seminal result of Savage [1972]. Here, instead of offering the agent a pool of lotteries between outcomes, the agent may observe a range of signals that allows the econometrician to pin down his initial belief over the distribution of the states.

Additionally, the full range assumption is necessary for point identification of the agent’s beliefs. If there is an interval on the positive real line for which there is no signal \( x \) whose likelihood ratio \( \gamma(x) \) lies within the interval (the range of \( \gamma(x) \) is not dense on the positive real line), then there will be two different beliefs \( p \) and \( p' \) which will be observationally equivalent. For any such interval, there is a pair or priors that can only be differentiated by a signal whose likelihood ratio falls within that interval.

In the following sections, we show that when the assumption of full range is relaxed the econometrician can still partially identify the priors of the agent.
2.2.1 Partial identification without full range

We now weaken the assumption of full range of information and consider partial identification when the range of the likelihood ratio is limited. Intuitively, this may happen when there is no information that is strong enough to either convince very pessimistic agents to invest or keep very optimistic agents from investing. Specifically, for the given information set, the range of the likelihood function is bounded, either above, away from zero, or both.

$$\text{Range}(\gamma(x)) \subset (\alpha, \bar{\alpha}) \subseteq \mathbb{R}_+$$

Within these bounds, we still assume that information is sufficiently dense as in the full range assumption from the previous section. In order for the econometrician to know the range of prior beliefs that can be identified, she must have information about the agent’s utility function. It is sufficient for her to know the value of $\bar{d}$, which determines the decision rule of the agent.

**Assumption 2.2.3.** The econometrician knows $\bar{d}$.

With this assumption, the econometrician can point identify prior beliefs of an agent who can be persuaded by available information. The beliefs of agents who cannot be persuaded by any of the available information will be partially identified.

**Proposition 2.2.2.** All beliefs of the agent for which $p(h)$ belongs to the interval

$$\left( \frac{\bar{d}}{\bar{d} + \alpha (1 - \bar{d})}, \frac{\bar{d}}{\bar{d} + \alpha (1 - \bar{d})} \right)$$

are point identified. If the belief is such that $p(h)$ is the right of the interval, then

$$H[p] = \left[ \frac{\bar{d}}{\bar{d} + \alpha (1 - \bar{d})}, 1 \right],$$

and if it is to the left of the interval

$$H[p] = \left[ 0, \frac{\bar{d}}{\bar{d} + \alpha (1 - \bar{d})} \right].$$

49
This proof follows the steps of Proposition 1 for the beliefs contained within the point
identified interval. If the agent’s beliefs are such that his decision does not change, his
beliefs can still be partially identified. For example, if the agent always decides to invest,
then his beliefs are such that \( p(h) \) is observationally equivalent to 1.

Note that without Assumption 3, or any knowledge about the agent’s utility function,
\( \bar{d} \) can take on any value between 0 and 1. In this case, the econometrician would not
know which beliefs are point identified and which are just partially identified. The region
of priors that are point identified is determined by \( \bar{d} \).

### 2.2.2 Partial identification with finite information

Point identification, even over a small region of the agent’s priors, requires an information
set that is infinite. Nevertheless, with assumptions 2.1 and 2.3, we can partially identify
the prior beliefs of the agent with a finite information set. The tightness of these bounds
will depend on how dense and how persuasive these observations are.

We now define \( \gamma^*(p) \) as the likelihood ratio that would make the agent indifferent
between investing and not investing given beliefs \( p \).

\[
p(h) = \frac{\bar{d}}{d + \gamma^*(p)(1 - \bar{d})} \Rightarrow \gamma^*(p) = \frac{\bar{d}(1 - p(h))}{1 - \bar{d}}
\]

Note that the agent will decide to invest after receiving signal \( x \) if and only if \( \gamma(x) > \gamma^*(p) \).
Using this, we can impose bounds on \( p(h) \).

Assume we have \( M \) possible signals given the information set \( X \), \( |X| = M \). We
can order the \( M \) potential observations according to the size of their likelihood ratio,
i.e. \( \gamma(x_{m-1}) \leq \gamma(x_m) \), for all \( 1 \leq m \leq M \). Then let \( \bar{\gamma} \) be the smallest \( m \) for which the
agent decides to invest after receiving signal \( x_{\bar{\gamma}} \). This would imply (for \( \bar{\gamma} > 1 \) that
\( \gamma(x_{\bar{\gamma}}) > \gamma^*(p) > \gamma(x_{\bar{\gamma}-1}) \) and the identification region is given by

\[
H_M[p(h)] = \left[ \frac{\bar{d}}{d + \gamma(x_{\bar{\gamma}})(1 - \bar{d})}, \frac{\bar{d}}{d + \gamma(x_{\bar{\gamma}-1})(1 - \bar{d})} \right].
\]  
(2.6)
If $\tilde{m} = 1$, then the agent always invests, and the identification region is

$$H_M[p_h] = \left[ \frac{\overline{d}}{\overline{d} + \gamma(x_1)(1-\overline{d})}, 1 \right].$$

On the other hand, if the agent never decides to invest, then the identification region is

$$H_M[p_h] = \left[ 0, \frac{\overline{d}}{\overline{d} + \gamma(x_M)(1-\overline{d})} \right].$$

### 2.2.3 Multiple States

We now consider a situation where the agent is making a binary decision, but there are more than two possible states of the world. This allows us to capture a situation where the agent has multiple reasons for why he would, or would not, want to invest. The decision rule is now a function of the agent’s posterior beliefs, $p(s|x)$, and gain from investing, $\tilde{v}(s)$, for each state $s \in \{1, \ldots, S\}$. From Bayes’ rule, these posterior beliefs are given by

$$p(s|x) = \frac{f(x|s)p(s)}{\sum_{s' \in S} f(x|s')p(s')} \text{ for all } s \in S. \quad (2.7)$$

Without loss, we order the states according to the gain in deciding to invest such that:

$$\tilde{v}(S) > \tilde{v}(S-1) > \cdots > \tilde{v}(1).$$

As in the two state model, we require the agent to prefer investing in at least one state and to prefer not investing in at least one different state.

**Assumption 2.2.4.** There exists a state $s^*$ such that $\tilde{v}(s^*) > 0 > \tilde{v}(s^* - 1)$.

For a given signal $x$, the agent will decide to invest exactly when the expected gain from investing is positive. Thus the agent invests when

$$E_{p(s|x)}[\tilde{v}(s)|x] = \sum_{s \in S} \tilde{v}(s)p(s|x) = \sum_{s \in S} \tilde{v}(s)p(s)f(x|s) \geq 0. \quad (2.8)$$

In order for the decision rule to be non-trivial, we assume that the agent is uncertain about whether the state of the world is good for investment. In this case, we say the agent has a prior belief that may be identified by the econometrician for a given information set.
An agent who is certain that the state is good, or certain that it is bad, will always make the same decision, and the econometrician will not be able to determine why this decision is being made.

**Definition 2.2.2.** A prior \( p \in \Delta(\mathbb{R}^5) \) is relevant if \( p(r) > 0 \) and \( p(t) > 0 \) for some states \( r < s^* \leq t \). Let the set of relevant priors be denoted by \( P \subset \Delta(\mathbb{R}^5) \).

For the econometrician to identify the beliefs of the agent in this broader setting, there must be information that not only separates beliefs by how optimistic they are, but also tell us why the agent is optimistic or pessimistic. The full range assumption which gives the econometrician sufficient information for identification is similar to the corresponding assumption when there are only two states. For each pair of states where one state is good for investing, and the other bad, we must have information that can sway both optimistic and pessimistic individuals, and be dense enough to separate beliefs that are similar, in this particular dimension. This assumption is formalized below.

**Assumption 2.2.5.** For each \( y \in \mathbb{R}_+, \varepsilon > 0, r < s^* \leq t \) there is an \( x \in X \), such that \( f(x|s) = 0 \) for all \( s \neq r, t \) and \( f(x|t)/f(x|r) = \gamma_x(x) \in B_\varepsilon(y) \).

The following lemma shows that the econometrician can point identify the priors of the agent when information satisfies full range and we can find two states, one good for investing, where the agents have different beliefs. The argument is similar to the proof of identification when there are two states. The proof is relegated to the appendix.

**Lemma 2.2.3.** For any \( p, p' \in P \), if there are two states \( r < s^* \leq t \) such that \( p'(t) \) and \( p'(r) \) are non zero and \( \frac{p(t)}{p'(t)} \neq \frac{p(r)}{p'(r)} \), then under the assumption of full range in multiple states, \( p' \not\in H[p] \).

To show identification of the set of relevant priors, it is now enough to show that any two distinct, relevant priors have the stated property of Lemma 1. Again the proof is in the appendix.
Theorem 2.2.4. The set of relevant priors is point identified if $X$ satisfies full range in multiple states.

While the full range in multiple states assumption appears strong, it cannot be weakened significantly. To see this, consider two priors that are indexed by 3 parameters, $k_1 > 1$, $k_2 > 1$, $n \geq 3$. Let $p(s; k_1, k_2, n) = 1 - \frac{1}{(k_1 + 1)n} - \frac{1}{k_2n}$, $p(r; k_1, k_2, n) = \frac{1}{(k_1 + 1)n}$, $p(t; k_1, k_2, n) = \frac{1}{k_2n}$, and $p(s'; k_1, k_2, n) = 0$ for all $s' \neq r, s, t$, where $r < s^*$ and $t \geq s^*$. Let $p'(s; k_1, k_2, n) = 1 - \frac{1}{k_1n} - \frac{1}{(k_1 + 1)n}$, $p'(r; k_1, k_2, n) = \frac{1}{k_1n}$, $p'(t; k_1, k_2, n) = \frac{1}{(k_1 + 1)n}$, and $p'(s'; k_1, k_2, n) = 0$ for all $s' \neq r, s, t$, where $r < s^*$ and $t \geq s^*$. Note that a decision maker with a prior of $p$ is more optimistic than one with a prior of $p'$ for all $k_1 > 1, k_2 > 1$, and $n \geq 3$. In order for $p'(k_1, k_2, n) \not\in H[p(k_1, k_2, n)]$, there must be an $x \in X$ such that the following inequalities hold.

\[
0 \leq f(x|s) \left[ \frac{(k_1 + 1)k_2n - (k_1 + 1) - k_2}{(k_1 + 1)k_2n} \right] \bar{v}(s) + f(x|r) \frac{1}{(k_1 + 1)n} \bar{v}(r) + f(x|t) \frac{1}{k_2n} \bar{v}(t) \tag{2.9}
\]

\[
0 > f(x|s) \left[ \frac{k_1(k_2 + 1)n - k_1 - (k_2 + 1)}{k_1(k_2 + 1)n} \right] \bar{v}(s) + f(x|r) \frac{1}{k_1n} \bar{v}(r) + f(x|t) \frac{1}{(k_2 + 1)n} \bar{v}(t) \tag{2.10}
\]

If $s \geq s^*$, then (10) implies that

\[
\gamma_{sr}(x) < \left[ \frac{k_2 + 1}{k_1(k_2 + 1)n - k_1 - k_2 - 1} \right] \frac{-\bar{v}(r)}{\bar{v}(s)}. \]

For large $n$, both decision makers are almost certain that they are in a state where investing is the optimal choice. In order to differentiate between the two, they must receive a piece of information that almost never occurs in this state. Condition (10) implies that the more pessimistic player must decide not to invest. Therefore $\gamma_{sr}(x) \to 0$ as $n \to \infty$.

Also, if $\gamma_{sr}(x) = 0$. Then condition (9) and (10) combine to require a piece of information for which the optimistic player invests, while the pessimistic player does not

\[
\frac{-\bar{v}(r)}{\bar{v}(t)} \frac{k_2}{(k_1 + 1)} < \gamma_{sr}(x) < \frac{-\bar{v}(r)}{\bar{v}(t)} \frac{k_2 + 1}{k_1}. \]

53
Therefore $\gamma_{sr}(x)$ must have full range while $\gamma_{sr}(x) = 0$. This argument can be made for each triple of states, $(s,r,t)$, where $r < s^*$ and $t \geq s^*$.

### 2.3 Identification of priors

For many economic applications, the idea that we observe an individual making the same decision repeatedly maybe of little practical purpose. In this section, we depart from the idea of identifying fixed, steady state beliefs and focus on the situation where there is only one move of nature that describes the state of the world and a series of signals about this underlying state. Provided that individuals have strictly positive priors and that we understand how they update their beliefs, there is a one to one mapping between priors and posteriors.\(^6\) Thus although we will focus on estimating priors, both objects are equivalent for predicting the agent’s behavior.

We return to the binary world and binary decision framework. Now the agent observes a sequence of signals, $X_M = \{x\}_{m=1}^M$, and after each signal he updates his posteriors and makes a decision, but he does not observe payoffs. In our investment problem, this assumption would correspond to the idea that the agent observes $M$ signals and makes $M$ subsequent decisions, before the state of the economy is revealed. We assume that the signals are independent, conditional on the state of the world.

The econometrician observes the sequence of signals $X_M$ and the sequence of decisions $D_M = \{d\}_{m=1}^M$. From these observations, she creates the sharpest bounds on the agent’s prior. The sharpest bounds constitute the smallest identification region for the given sequence of observations. After each signal, the agent will invest exactly when the expected value of investment is positive, given his current posterior beliefs:

\(^6\)Acemoglu et al. [2009] showed that if individuals disagree or are uncertain about the interpretation of the signals, there will be no asymptotic learning and no asymptotic agreement. Rather, even if individuals have the same initial prior and observe the same infinite sequence of signals, individuals can reach different posterior beliefs.
\[ E_{p(s|X_m)}[\hat{v}(s)|X_m] \geq 0. \] This holds whenever \( p(h|X_m) \geq \bar{d}. \)

In order to highlight the one-to-one relationship between posteriors and priors, we take the log of the ratio of the posterior. Additionally, we can separate the effect that each signal has on the posterior of the agent. The assumption that the signals are conditionally independent given the state allows us to describe this relationship in the following equation:

\[
\ln \left( \frac{p(h|X_M)}{p(\ell|X_M)} \right) = \ln \left( \frac{f(x_1|h) \cdots f(x_M|h)p(h)}{f(x_1|\ell) \cdots f(x_M|\ell)p(\ell)} \right) = \ln \left( \prod_{m=1}^{M} \gamma(x_m) \left( \frac{p(h)}{p(\ell)} \right) \right) = \sum_{m=1}^{M} \ln(\gamma(x_m)) + \ln \left( \frac{p(h)}{p(\ell)} \right) \tag{2.11}
\]

The following proposition provides the sharpest bounds the econometrician can impose on the agent for a given sequence of signals and decisions.

**Proposition 2.3.1.** The sharpest bounds in the region of identification of priors are given by

\[
H \left[ \ln \left( \frac{p(h)}{p(\ell)} \right) \right] = \left[ \max_{1 \leq m \leq M} \ln \left( \frac{\bar{d}}{1-\bar{d}} \right) - \sum_{m=1}^{M} \ln(\gamma(x_m)) \right] - \left[ \sum_{m=1}^{M} \ln(\gamma(x_m)) \right].
\]

**Proof.** In every period \( j \), the agent will decide to invest if and only if

\[
\sum_{m=1}^{j} \ln(\gamma(x_m)) + \ln \left( \frac{p(h)}{p(\ell)} \right) \geq \bar{d}.
\]

Thus whenever the agent chooses to invest, given observation pair \( \{x\}_{m=1}^{j}, d_j \), we can impose a lower bound for \( \ln(p(h)/p(\ell)) \). The sharpest lower bound is derived from the period where the agent has the lowest posterior but still decides to invest. This is exactly

\[
\max_{1 \leq m \leq M} \ln \left( \frac{\bar{d}}{1-\bar{d}} \right) - \sum_{m=1}^{M} \ln(\gamma(x_m)).
\]

Similarly the sharpest upper bound is formed from the period where the agent has the highest posteriors given that he decides not to invest. \( \square \)
Figure 2.1: Constructing the bounds on priors.
Figure 1 depicts an example of the construction of these bounds. The econometrician knows \( \overline{d} \), which is denoted as the horizontal line, and she also knows the summation of the log likelihood ratio of the sequence of signals. On the figure, this is the vertical distance between the log of the initial beliefs and the log of the current beliefs. In this example, the upper bound is derived from observation 4, which is the highest posterior for which the agent chooses not to invest. The lower bound is constructed from observation 7. We construct each bound by subtracting this vertical distance from the \( \overline{d} \) in each key observation. After finding the bounds of \( \ln(p(h)/p(\ell)) \) we can use them to derive bounds for \( p(h) \) and \( p(\ell) \).

There is no reason to believe that the inference about the prior will necessarily improve asymptotically. However, since we are working with conditionally independent signals, the posterior of the agent approaches the true state of the world, so identification becomes less important.

From this example it is clear that all observations can be important in forming bounds on the priors of the agent. The econometrician should be concerned with all the past decisions, not because the present world mechanically depends on the past, but because past observations tighten the bounds on the agent’s beliefs.

### 2.4 Concluding remarks

In this chapter we directed our attention to a classical question in economics: what can we infer about an agent’s beliefs about the state of the world if we observe his decisions? We believe that there is a growing class of problems where economists are interested in an agent’s beliefs and are comfortable making assumptions about his utility function, or at least about his decision rule. We show that even in binary decision making with multiple states, agents beliefs can be precisely identified if we observe a range of signals that are sufficiently rich. We hope that these positive theoretical results will lead to further
research agendas, specifically in experimental economics and structural economics where identifying beliefs is the main goal.
Appendix

Lemma 1

For any $p, p' \in P$, if there are two states $r < s^* \leq t$ such that $p'(t)$ and $p'(r)$ are non zero and $\frac{p'(r)}{p'(t)} \neq \frac{p'(r)}{p'(r)}$, then under the assumption of full range in multiple states, $p' \not\in H[p]$.

Proof. We take $\frac{p'(r)}{p'(t)} > \frac{p'(t)}{p'(r)}$ (the argument is symmetric for the case where $\frac{p'(r)}{p'(t)} < \frac{p'(t)}{p'(r)}$). Then,

$$\frac{p'(r)}{p'(t)} > \frac{p(t)}{p'(t)} \Rightarrow \frac{\tilde{v}(r)p(r)}{\tilde{v}(r)p'(r)} > \frac{\tilde{v}(t)p(t)}{\tilde{v}(t)p'(t)}$$

$$\Rightarrow \frac{\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)} < \frac{\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)} \Rightarrow \frac{-\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)} > \frac{-\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)}$$

For $p' \not\in H[p]$, we must have that $d(p, x) \neq d(p', x)$ for some $x \in X$. Consider $X_{rt} = \{x \in X : f(x|s) = 0 \text{ for } s \neq r, t\}$. Then for $x \in X_{rt}$, $d(p, x) \neq d(p', x)$ exactly when

$$\tilde{v}(t)p(t)f(x|t) + \tilde{v}(r)p(r)f(x|r) < 0 \text{ and } \tilde{v}(t)p'(t)f(x|t) + \tilde{v}(r)p'(r)f(x|r) \geq 0.$$  

Dividing by $f(x|r)$

$$\tilde{v}(r)p(t)\gamma_{rt}(x) + \tilde{v}(r)p(r) < 0 \text{ and } 0 \leq \tilde{v}(t)p'(t)\gamma_{rt}(x) + \tilde{v}(r)p'(r)$$

Combining the inequalities,

$$\frac{-\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)} \leq \gamma_{rt}(x) < \frac{-\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)}$$

Now let

$$y = \frac{1}{2} \left( \frac{-\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)} + \frac{-\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)} \right), \text{ and } \epsilon = \frac{1}{3} \left( \frac{-\tilde{v}(r)p(r)}{\tilde{v}(t)p(t)} - \frac{-\tilde{v}(r)p'(r)}{\tilde{v}(t)p'(t)} \right)$$

Then there is an $x \in X_{rt} \subseteq X$ such that $\gamma_{rt}(x) \in B_{\epsilon}(y)$ and for this $x$, $d(p, x) \neq d(p', x)$ and therefore $p' \not\in H[p]$. \hfill \Box
Theorem 1

The set of relevant priors is point identified if $X$ satisfies full range in multiple states.

Proof. Let $p, p' \in P$ where $p \neq p'$. Then there is an $s \in S$, where $p(s) \neq p'(s)$. Without loss of generality, we take $0 \leq p(s) < p'(s)$.

If $s < s^*$ then there is a state $t \geq s^*$ where $p'(t) > 0$, which implies $\frac{p(s)}{p'(s)}$ and $\frac{p(t)}{p'(t)}$ are non-negative numbers. If $\frac{p(s)}{p'(s)} \neq \frac{p(t)}{p'(t)}$, then by Lemma 2.2 we know that $p' \not\in H[p]$. If $\frac{p(s)}{p'(s)} = \frac{p(t)}{p'(t)}$, then since $\sum_{i=1}^{S} p(i) = \sum_{i=1}^{S} p'(i) = 1$, there must be an $s'$ where $0 \leq p'(s') < p(s')$. If $s' < s^*$ then $\frac{p(s')}{p'(s')} \neq \frac{p(t)}{p'(t)}$, and we can use Lemma 2.2 to show that $p' \not\in H[p]$. If $s' \geq s^*$ then $\frac{p(s)}{p'(s)} \neq \frac{p(s')}{p'(s')}$, and we can again use Lemma 2.2.

Now if $s \geq s^*$, then there is a state $r < s^*$ where $p'(r) > 0$. We can now repeat the argument above to show that $p' \not\in H[p]$.

Since $p$ and $p'$ were arbitrary, then $H[p]$ is a singleton for all $p \in P$, and $P$ is point identified. \qed
CHAPTER 3

Playing with Altruism

3.1 Introduction

"How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. ...That we often derive sorrow from the sorrow of others, is a matter of fact too obvious to require any instances to prove it; for this sentiment, like all the other original passions of human nature, is by no means confined to the virtuous and humane, though they perhaps may feel it with the most exquisite sensibility. The greatest ruffian, the most hardened violator of the laws of society, is not altogether without it."

Adam Smith, TMS, Book I, Chapter I

Since the beginning of the study of economics, altruism has been understood as an important force in human nature. Smith [2010] argued that not only the virtuous but all individuals possess it to some degree. Nonetheless, the analysis of the implications of altruism in economics are mostly restricted to development economics and household settings and there are few works attempt to analyze its impact outside the family environment. The probable explanation for this restriction is the assumption that only in some restricted environments, agents are altruistic enough to actually matter.
The goal of this chapter is to show that even small levels of altruism lead to inefficient decision-making and manipulation. More specifically, I will focus on the effect of endowments in individual’s utilities. Usually, in economic settings, the higher the initial endowment, the better the individual will be. Indeed, in Nash bargaining developed by Nash [1953] and in Sequential Bargaining models as in Rubinstein [1982], the higher the disagreement point, the higher the output from the bargain will be.

Despite these core results in economics, it is not clear whether in every situation the utility monotonically increases with the endowment. I intend to show that this is not true. Let us consider an example when part of the utility of an agent comes from a decision of another altruistic agent, in the form of a gift. If this Decision Maker takes into account the endowment of the agent in the optimal allocation of the gift then the monotonicity condition can be violated.

This chapter will provide a theoretical framework for sequential decision making with altruism. A couple applications are important to sharpen our intuition. For instance, when the decision maker is a manager who has to choose which employee to fire, an altruistic manager might choose to keep the less qualified, instead of the more qualified, who has higher endowments and better outside option, even in the case when he cares equally for both of them.

In the second section of the chapter, I develop the divisible goods model. There is an altruist decision maker who decides how to split gift to two other agents. The agents derive utility from the consumption of his own endowment and from the consumption of the gift that he receives.

For simplicity, I assume that it is efficient to simply allocate the good to agent 1. However, because the Decision Maker is altruist, he will split the gift unequally between the two. Moreover, I show that if endowments and consumption are complementary goods, then the utility will be strictly increasing in the endowments. However, if they are substi-
tute, it is possible that a lower endowment diminishes directly the utility, but it also forces the altruist decision maker to give the agent more consumption good and, because they are substitutes, this effect overcomes the direct effect.

Although this situation seems extreme with divisible goods, it can easily arise with indivisible gifts. In order to illustrate this effect, in section 3, I solve a model that is similar to the previous chapter but the good is indivisible. Assuming extreme substitutability, I show that the value function is not strictly increasing in the total endowments. The intuition is similar to what was discussed before, there is a unique value of endowments, $d^*$, at which the decision maker will change his mind and allocate the good to the inefficient agents. Therefore, whenever this agent has an endowment slightly lower than this threshold, he will be strictly better than slightly above.

Following this intuition, in section 4, I develop a simplified model sequential model where firstly agents invest in their endowments and in the second period, the altruist decision maker sees the endowments and decides to whom he wants to allocate the indivisible good. I show that there is always a mass of agents who want to invest less than optimally in their endowments and receive the gift. Moreover, the mass of these manipulative players is strictly increasing in the level of altruism.

The model described in rest of this chapter encompasses many possible applications. It also contributes to some debates in development economics. Becker [1974] states the "Rotter Kid theorem" according to which even if kids were completely selfish they still would have incentives to take actions that maximize the household allocation. Bergstrom [1989] details the restrictions of this theorem, finds sufficient and necessary conditions and shows that it holds for only a small set of utilities. The experimental literature often estimates that fathers are not altruistic enough towards their kids and economists have given recommendations that provision goods should be given to the mothers. However, most experiments are not made in life or death situations, thus, we have reasons to suspect that in such situations, when the decision is extremely important to his heirs, fathers will
act altruistically.

Lindbeck and Weibull [1988], a paper that is the most similar to this work, studies altruism and time consistency and shed light in the Prodigal Son parable. They argue that, if heritage of the parents depends on the endowments of their heirs, these, anticipating their decisions, might not have the incentives to take the actions that maximize their allocation without heritage.

In all these cases, despite the literature in experimental economics and in psychology stating the people are altruistic in several situations\(^1\), the unit of observation was the household. Perhaps the reason to focus on the household environment is because the literature usually considers that if altruism exists, it should not be big enough matter in outside home situations.

The rest of this chapter is organized as follows. The next section will show that inefficiencies with altruism are present in a model with divisible goods. In the third section I will work on a model of indivisible goods and show that this inefficiency can arise, no matter how altruistic the decision maker is. In section 4 characterizes the measure of manipulative agents and section 5 concludes.

### 3.2 Divisible Goods Model

In order to start our analysis, let us consider a simplistic divisible goods model. I hope to depict that some of the features of decision making and altruism are already present in this model.

We have one Decision-Maker and two other individuals. The DM is the only altruistic and cares about both individuals equally. The good is divisible and he has to choose how to split it among them. Hence, his utility is:

\(^1\)See Krebs [1970] and Levine [1998]
\[ U^{DM}(Y, U^1, U^2) = Y + \alpha U^1 + \alpha U^2 \]

Where \( \alpha \) is the altruism parameter. The other two individuals utilities are

\[ U^i(x_i, d_i) \]

strictly increasing and concave in both arguments for \( i = 1, 2 \). Where \( x_i \) represents the good allocated to individual \( i \) and \( d_i \) represents the endowment he possess. The DM allocates the gift \( x \in \mathbb{R}^2_+ \) between these individuals such that \( x_1 + x_2 = 1 \).

I also assume that the DM only obtains direct utility from given the good to individual 1. That is \( Y = x_1 \). This assumption has different interpretations in different models, for example we could simply consider that individual 1 is more productive. I only want to impose that without altruism the economic rationality would led all the good to be given to 1. Therefore the DM problem becomes:

\[
\max_{x_1} x_1 + \alpha U^1(x_1, d_1) + \alpha U^2(1 - x_1, d_2)
\]

Therefore his FOC implies:

\[ 1 + \alpha U^1_x = \alpha U^2_x \]

Where subscripts denote partial derivatives and stars denote optimal decisions.

Now the relevant question for us is how does this decision rule change with a change in the endowment of individual 2. First notice that \( \frac{\partial x^*_2}{\partial d_2} = -\frac{\partial x^*_1}{\partial d_2} \). Using implicit function theorem we obtain:

\[
\frac{\partial x^*_2}{\partial d_2} = -\frac{U^2_{sd_2}}{U^1_{xx} + U^2_{xx}}
\]
Since both utilities are concave, the numerator is always negative, therefore the relationship between endowments and gifts depends only in the cross derivative\(^2\). That is:

\[
\frac{\partial x^*_2}{\partial d_2} > 0 \iff U^2_{xd_2} > 0
\]

The condition above is straightforward, the gift the individual receives will increase with his endowments if and only if his marginal utility to the gift increases with higher endowments.

But we are interested in a more general case. When does the utility of the agent 2 increases with his initial endowment? One channel is direct, the utility is increasing in endowment. The other is indirect, the share of the gift might or might not augment with endowments. Since the individuals anticipate the decision rule of the DM we can write his utility as:

\[
U^2(x_2, d_2) = U^2(x^*_2(d_2), d_2) = g(d_2)
\]

We would like \(g(d_2)\) to be strictly increasing in \(d_2\), so that agents could do everything in their power to maximize their endowments in other aspects of their life. Nevertheless the anticipation of the gift given by the DM may lead agents to manipulate their endowments in order to maximize the total utility. The effect of the endowment in the utility can be summarized as:

\[
g'(d_2) = U^2_x \cdot \frac{\partial x^*_2}{\partial d_2} + U^2_{d_2}
\]

Therefore the utility will be increasing in the endowment if and only if:

\[
g'(d_2) > 0 \iff U^2_x U^2_{xd_2} > U^2_{d_2} (U^1_{xx} + U^2_{xx})
\]

\(^2\)The fact that it does not depend on \(\alpha\) comes from the linearity of the relation \(Y = x_i\) and has no special economics meaning.
Notice that the right hand of the inequality is always negative, hence $U_{i,d_2}^2 > 0$ is a sufficient condition for the inequality to hold. Indeed, most of the literature tend to focus on the case where the above equality does hold. Because the cross derivative is positive or when it’s negative it is too small to compensate the direct gain of more endowment.

In the rest of this chapter, I show that indivisible goods can aggravate these time inconsistencies and increase the inefficiencies that arise with altruism. Perhaps the strongest signal that this should be a concern is that these inefficiencies will occur no matter how altruistic the Decision Maker is.

3.3 Indivisible Goods Model

In this section, just as in the divisible goods model, the Decision Maker is equally altruistic among individuals and his utility function has the form of:

$$U^{DM}(Y,U^1,U^2) = Y + \alpha U^1 + \alpha U^2$$

But now, differently from the previous model, he allocates an indivisible good to one of the individuals, there is a gift $x \in X = \{(0,1); (1,0)\}$.

We also want to tilt the decision towards individual one, hence DM only obtains direct utility from allocating the good to individual one. Hence

$$Y = \mathbb{1}[x = (1,0)]$$

I will impose two assumptions in the individuals’ utility functions

**Assumption 3.3.1.** Individuals utility: $U^i(x_i,d_i)$ is continuous, strictly increasing and concave in both arguments.
**Assumption 3.3.2.** Extreme Substitutability

\[
\lim_{d_i \to 0} [U^i (1, d_i) - U^i (0, d_i)] = \infty
\]

\[
\lim_{d_i \to \infty} [U^i (1, d_i) - U^i (0, d_i)] = b > 0
\]

This is a restrictive assumption, but it fits well our interpretation that \(d_i\) is fundamentally an endowment. The intuition is that as the endowment approaches zero, receiving the gift becomes crucial to individuals and even when individuals are sufficiently wealthy, this gift still matters. Figure 1 depicts the utility function of an agent with the gift, \(U_2(1, d_2)\) and without it, \(U_2(0, d_2)\).
Is easy to see that the decision maker will allocate the good to individual 1 if and only if:

\[ Y((1,0)) + \alpha U^1(1,d_1) + \alpha U^2(0,d_2) \geq Y((0,1)) + \alpha U^1(0,d_1) + \alpha U^2(1,d_2) \]

Rearranging the terms and using the definition of \( Y \) this condition becomes:

\[ 1 \geq \alpha \left[ (U^2(1,d_2) - U^2(0,d_2)) - (U^1(1,d_1) - U^1(0,d_1)) \right] \]

Evidently, for \( \alpha \) sufficient low the decision maker will never allocate the gift to individual 2 and if \( d_2 < d_1 \) for a level of altruism sufficiently high, he will.

Again, individual 2 knows that this is the allocation rule the DM follows and will anticipate this decision. Hence we can write:

\[ U^2(x_2,d_2) = U^2(x^*_2(d_2),d_2) = g(d_2) \]

In this setting we can prove Theorem 1, that form every \( \alpha > 0 \), the function \( g(d_2) \) is non-monotonic. I leave the details of the proof for the appendix but the outline follows the steps of the above discussion. There is one unique \( d^* \) which will make the decision maker change his mind, so the value function will drop just after that level and will increase elsewhere.

**Theorem 3.3.1.** \( g(d_2) \) is non-monotonic. For every \( \alpha > 0 \) the function \( g(d_2) \) is non-monotonic

Figure 2 provides an illustration of the function \( g(d_2) \) that takes into account the problem that the decision maker will have to face.

If we understand the endowments not only as an exogenous factor that factors into the utilities but as a result of series of previous actions that they have taken to maximize
their well-being, we can understand why non-monotonicity becomes a problem. If endowments cannot be verified, individuals always have the incentive to misreport it to the Decision Maker and try to gain the gift. However even if endowments are verifiable, for every level of altruism, a mass of agents will take suboptimal decisions in previous periods, just so that they can manipulate the Decision Maker’s allocation. In the next section, I will properly define and characterize this mass of manipulative agents.

### 3.4 Manipulative Agents

In this section, I will focus on characterizing the measure of agents who have incentives to manipulate their endowments in order to obtain the gift from the decision maker. Let us first define this measure $d'$ as the of level of endowment that would make agent 2
indifferent between having $d'$ and no gift, and have the gift a the point that the decision maker changes his mind, that is:

$$U^2(0,d') = U^2(1,d^*)$$

Since $U^2(.,d_2)$ is increasing in both arguments, we have that $d' > d^*$. Notice that $d'$ is function of $d^*$ and therefore we can define the mass of manipulative agents as $\mu(d^*)$:

$$\mu(d^*) = d' - d^*$$

Figure 3.3: Manipulative Agents

In order to better describe the mass of manipulative agents, we need to impose one more assumption in the utility function. Instead of only use extreme substitutability we will assume:
Assumption 3.4.1. Monotone Substitutability

\[ [U^i(1,d_i) - U^i(0,d_i)] \text{ is decreasing in } d_i \]

Using this assumption and the previous definition, we can proof the following theorem.

Theorem 3.4.1. Characterization of the mass of manipulative agents. \( \mu(\alpha^*) \) is increasing in the level of altruism \( \alpha \) and in \( d_1 \)

This result is intuitive: the more altruistic the decision maker is, the more agents will be tempted to manipulate their outside options and obtain the gift. Moreover, the higher the endowment of individual 1, the higher will be the measure of individuals type 2 who have incentives to do so. In figure 3 we can see a characterization of the above result describing the mass of manipulative agents.
3.5 Conclusion

Often, in economic settings, we assume that the higher the outside option the better the person will be. However, this is not necessarily the case when one of the players is altruistic. I construct a model where an altruistic Decision Maker allocates a gift to another player and showed that the utility of the player who receives it can decrease with his endowment.

This will be the case with indivisible goods, in that framework, there is always a level of the endowment which makes the decision maker change his mind about to whom allocate the gift, this creates a natural threshold for the initial endowment. Endowments higher than this threshold, he will give the gift to another player. In a game wherein the first stage agents invest in their endowments and in the second stage the decision maker allocates the gift to one of the agents, there will be a mass of agents who have incentives to manipulate their endowments in order to get the gift. I characterize this mass of agents and show that they are increasing with the amount of altruism.
3.6 Mathematical Appendix

Proof of Theorem 1. Fix $\alpha > 0$ and $d_1 > 0$.

Let $U^1(1,d_1) - U^1(0,d_1) \equiv c$. By assumption $1 \ c > 0$ Individual 1 will receive the gift iff:

$$\frac{1}{\alpha} + c \geq U^2(1,d_2) - U^2(0,d_2)$$

By Ass 2: $\exists d_2$ such that this inequality does not hold and $\overline{d_2}$ such that it holds.

Since $U^2(x_2,d_2)$ is continuous function, $\exists d^*$ such that equality holds.

Take $\varepsilon > 0$, this implies that

$$g(d^* + \varepsilon) = U^2(0,d^* + \varepsilon) \text{ and } g(d^* - \varepsilon) = U^2(1,d^* - \varepsilon)$$

Since utility function is continuous and increasing, for $\varepsilon$ small enough

$$g(d^* - \varepsilon) > g(d^* + \varepsilon)$$

Everywhere else $g(d_2)$ is increasing. \hfill \square

Proof of Theorem 2. Using the definition of $d^*$:

$$\frac{1}{\alpha} + c = U^2(1,d^*) - U^2(0,d^*)$$

The LHS does not depend on $d_2$ while, by monotonic substitutability the RHS is decreasing in $d_2$. Since everything that decreases LHS increases $d^*$, $\alpha$ and, by monotonic substitutability, $d_1$ increase $d^*$.

Now we need to show that $\mu(d^*)$ is increasing with $d^*$.

Using the definition of $\mu(d^*) = d' - d^*$. Differentiating w.r.t $d^*$:
\[
\frac{\partial \mu (d^*)}{\partial d^*} = \frac{\partial d'}{\partial d^*} - 1
\]

Using the definition of \(d'\) that \(U^2 (0, d') = U^2 (1, d^*)\), we have:

\[
\frac{1}{\alpha} + c = U^2 (0, d') - U^2 (0, d^*)
\]

Differentiating w.r.t \(d^*\) we have:

\[
0 = U^2_{d_2} (0, d') \cdot \frac{\partial d'}{\partial d^*} - U^2_{d_2} (0, d^*)
\]

Rearranging the terms:

\[
\frac{\partial d'}{\partial d^*} = \frac{U^2_{d_2} (0, d^*)}{U^2_{d_2} (0, d')}
\]

Since \(U^2_{d_2} (0, d_2)\) is increasing and concave and \(d' > d^*\), we have that \(\frac{\partial d'}{\partial d^*} > 1\). Therefore

\[
\frac{\partial \mu (d^*)}{\partial d^*} > 0
\]
Bibliography


Itai Arieli and Manuel Mueller-Frank. Inferring beliefs from actions. *Available at SSRN 2208083,* 2014.


