Essays on time-inconsistency and revealed preference

Paweł Dziewulski

St Hilda's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2014

Agacie, za wszystko

Acknowledgements

I would like to thank my supervisor, John Quah, for his invaluable help during writing this thesis. I am grateful for his advice, his support, and the exorbitant amount of time spent on discussing, revising, and improving my work. I would like to thank him for his encouragement and guidance throughout the years which allowed to grow and become a far better economist than I have ever expected to be.

I would like to thank Łukasz Woźny for the support and encouragement at the early stage of my academic career. In particular, I am grateful for a certain conversation that took place 8 years ago, which convinced me to follow the path of scientific research. Without his guidance and the faith put in me, I would never be where I am today.

It would not have been possible to write this doctoral thesis without the help and support of the kind people around me. I am especially grateful to Łukasz Balbus, Alan Beggs, Dan Beary, Vince Crawford, Francis Dennig, Glenn Harrison, Paweł Gola, Faruk Gul, Jakub Growiec, Jean-Jacques Herings, Claudia Herresthal, Max Kwiek, Yusufcan Masatlioglu, Meg Meyer, Sujoy Mukerji, Drazen Prelec, Krzysztof Pytka, Collin Raymond, Kevin Reffett, Kota Saito, Koji Shirai, Bruno Strulovici, Tomasz Strzalecki, and Bassel Tarbush for many insightful comments and suggestions at various stages of my research.

Finally, I would like to thank my family for their support and encouragement in all my pursuits. Foremost, I am thankful to my wife, Agata, my muse, my rock, my love.

Abstract

This thesis concerns three important issues related to the problem of time-inconsistency in decision-making and revealed preference analysis.

The first chapter focuses on the welfare properties of equilibria in exchange economies with time-dependent preferences. We reintroduce the notion of time-consistent overall Pareto efficiency proposed by Herings and Rohde (2006) and show that, whenever the agents are sophisticated, any equilibrium allocation is efficient in this sense. Thereby, we present a version of the First Fundamental Welfare Theorem for this class of economies. Moreover, we present a social welfare function with maximisers that coincide with the efficient allocations and prove that every equilibrium can be represented by a solution to the social welfare optimisation problem.

In the second chapter we concentrate on the observable implications of various models of time-preference. We consider a framework in which subjects are asked to choose between pairs consisting of a monetary payment and a time-delay at which the payment is delivered. Given a finite set of observations, we are interested under what conditions the choices of an individual agent can be rationalised by a discounted utility function. We develop an axiomatic characterisation of time-preference with various forms of discounting, including weakly present-biased, quasi-hyperbolic, and exponential, and determine the testable restrictions for each specification. Moreover, we discuss possible identification issues that may arise in this class of tests.

Finally, in the third chapter, we discuss the testable restrictions for production technologies that exhibit complementarities. Suppose that we observe a finite number of choices of input factors made by a single firm, as well as the prices at which they were acquired. Under what conditions imposed on the set of observations is it possible to justify the decisions of the firm by profit-maximisation with production complementarities? In this chapter, we develop an axiomatic characterisation of such behaviour and provide an easy-to-apply test for the hypothesis which can be employed in an empirical analysis.

Preface

The questions we consider in this thesis lie in the intersection of general equilibrium theory, behavioural economics, and revealed preference analysis. Our discussion evolves around three main topics.

First of all, we are interested in the welfare properties of economies with time-dependent preferences. Suppose that, each period, every consumer is represented by a different *self*, whose preferences are defined over sequences of consumption bundles from the current period till the end of time. If a credible commitment device is not available, the current self needs to take into account the behaviour of his future incarnations while choosing his lifetime consumption, since any plan determined in the current period may be revised by one of his future selves. In this thesis, we determine properties of allocations that arise from trade between consumers characterised in the above manner. In particular, we are interested whether equilibrium outcomes are efficient in any sense, and how the dynamic structure of trade influences the final distribution of consumption.

Time-inconsistency of preferences plays a fundamental role in some of the major economic issues, including the problem of sustainable growth, the efficient use of depletable resources, or the question of optimal pension schemes. Understanding how present-bias and preference for immediate gratification affect decisions of individual agents as well as the aggregate behaviour of the economy is crucial for developing a sound economic policy that would address the issues mentioned above.

In the first chapter of the thesis we concentrate on the general properties of allocations that arise in this class of economies. We re-introduce the notion of *time-consistent overall Pareto efficiency*, defined by Herings and Rohde (2006), according to which an allocation is efficient if, at any time t, there exists no other feasible allocation which improves upon the original one with respect to the preferences of all the current and future selves from period t onwards. The definition takes into account two important factors. First of all, it requires that the allocation is Pareto efficient with respect to all the agents and all their different selves. Second of all, it imposes a form of time-consistency on the efficient allocations. This is to say that, as the time progresses and the initial selves are gradually excluded from the economy, the remaining incarnations cannot benefit from altering the efficient allocation.

The main result of the chapter shows that any general equilibrium allocation is efficient in the above sense. Hence, we present a version of the First Fundamental Welfare Theorem for economies with time-dependent preferences. The result is striking, as it shows that time-inconsistency of preferences is not a source of inefficiency as long as we are concerned with the welfare of all agents and their different selves. That is, any alteration of an equilibrium allocation negatively affects the well-being of at least some incarnations of some consumers. Therefore, any economic policy should be designed taking into account the full extent to which it affects the welfare of the agents, as there is no clear Pareto improvement which could make all the consumers and all their different selves better off.

In the second chapter of the thesis we remain in the area of inter-temporal decision making and time-dependent preferences, however, we focus on the individual behaviour of agents, rather than market outcomes. In this part of our discussion we focus on the testable restrictions for various models of inter-temporal choice. We are interested in conditions imposed on sets of observations that allow us to justify the observable choices of individual agents with a certain type of time-preference under the utility-maximisation hypothesis. Our main interest is focused on the class of discounted utility models with various specifications of the discounting function.

The discounted utility model plays a crucial role throughout the economic analysis and is widely accepted as a valid normative standard for public policies, as well as a descriptively accurate representation of the actual behaviour of economic agents. However, in the recent years an important question was raised concerning the form of the discounting function that reflects the actual time-preferences of consumers. In particular, alternative specifications of hyperbolic and quasi-hyperbolic discounting were proposed, which could explain various observations anomalous in the model of exponential discounting utility, formerly dominant in economics. In order to understand the actual behaviour of economic agents, we find it crucial to determine what are the observable implications of the above models and how can we differentiate between the various specifications of discounting, given a finite set of observations.

In our analysis we consider a framework in which agents are allowed to choose between pairs that consist of a monetary payment and a time-delay at which the payment is delivered. In each trial of the experiment an agent is offered a finite set of options from which he is allowed to choose a single element. Therefore, each observation consists of a set of feasible options from which the agent could choose and the choice made by the subject. Given that the observer is allowed to monitor only a finite number of repetitions of the experiment, we provide an axiomatic characterisation of various specifications of time-preference. In particular, we establish the verifiable implications of the separable, discounted utility model and present a simple test that can be applied to the real-life data.

Our analysis builds a bridge between the decision-theoretical approach to consumer choice, which characterises different forms of behaviour, and the experimental work, that elicits the tastes of agents from a finite list of observable choices. Our main objective is to fill the gap between these two areas of economic research. We provide an intuitive consistency condition that has to be satisfied by observable choices made by agents who are endowed with a certain type of time-preference. At the same time the condition allows us to construct an easy-to-apply test to verify or refute the model of inter-temporal decision making on the data set.

Our framework is especially relevant for empirical applications. There are numerous examples of experiments performed in the literature in which subjects were asked to choose between monetary payments delivered with various time-delays. Therefore, our analysis can be directly applied to observations from the experiments performed in the literature, in order to empirically determine the form of time-preference that is most likely to describe the behaviour of economic agents.

In the final part of the thesis, co-authored with John K.-H. Quah, we move away from questions related to time-preference and consumer choice in order to focus on the behaviour of firms and the testable implications for production with complementarities. Given a finite number of observations of input factors and prices at which they were acquired, we determine conditions imposed on the set of observations under which we can rationalise the decisions of the firm by the profit-maximising behaviour with production complementarities. We refer to the notion of complementarity introduced by Edgeworth, according to which two inputs are complements whenever an increase of one of the factors increases the marginal returns from the other one.

The importance of complementarities for modern manufacturing follows from the fundamental shift from mass production to a new pattern of manufacturing based on flexibility and economies of scope, that took place in the final decades of the twentieth century. This new paradigm relies on a system-wide and coordinated approach to production, where the "fit" of various attributes of technology plays a fundamental role. At the same time, the mathematical representation of complementarity in terms of supermodular functions, and the one of lattice programming techniques generally, provides a way of formalising the intuitive ideas of synergy and system effects and allows for a rigorous analysis of their implications. Given the importance of the notion of complementarity and its mathematical formalisation, it is crucial to determine whether the hypothesis provides any observable restrictions that could be tested.

The main purpose of our analysis is to provide an axiomatic characterisation of the profit-maximising behaviour with production complementarities. Additionally, we present an easy-to-apply test which allows us either to confirm or refute the hypothesis on data sets.

We consider our thesis to be aimed at a wide audience. On one hand, we answer significant theoretical questions regarding the welfare properties of general equilibrium economies, an axiomatic characterisation of consumer behaviour, and the fundamental features of technologies with complementarities. On the other hand, we discuss important issues related to the testable implications of the hypotheses in question and provide methods which would allow for their empirical verification. Furthermore, we consider topics that are related to various areas of economic research, including general equilibrium theory, macroeconomics, welfare economics, decision theory, behavioural economics, theory of the firm, revealed preference analysis, and experimental economics. Therefore, we hope that the work presented on the following pages will find a deep interest of different types of readers.

Contents

Chapter 1

Efficiency and representation of competitive equilibria in economies with time-dependent preferences

1.1 Introduction

Consider an exchange economy consisting of consumers endowed with time-dependent preferences. As in Strotz (1955), each period an agent is represented by a different *self* whose preferences are defined over sequences of consumption bundles from the current period till the end of time. Since tastes may differ across dates in an unrestricted manner, each consumer is characterised by a sequence of preference relations of all his subsequent selves. If a credible commitment device is not available, each period the current self needs to take into account the behaviour of his future incarnations while choosing his consumption, since any plan determined in the current period may be revised by one of his future selves. Assuming that agents are sophisticated, that is, they can correctly predict changes in the preferences of their subsequent incarnations, the demand is determined by a subgame perfect Nash equilibrium path of the game between different selves of an individual consumer.[1]

Suppose that we allow the agents to trade. Every period the current selves may exchange their rights to consumption inherited from the preceding period for a different consumption plan which is affordable, given the current prices. In this chapter we characterise the welfare properties of equilibrium allocations arising in this class of economies. We describe a notion of efficiency and present conditions under which

[1]Our use of the term *sophisticated* follows, e.g., Pollak (1968).

every equilibrium allocation is efficient in the sense defined. Moreover, we construct a social welfare function with maximisers that coincide with the efficient allocations. The two results allow us to present a method of representing equilibria by a solution to a social planner's optimisation problem.

In our work we reintroduce the notion of *time-consistent overall Pareto efficiency* proposed by Herings and Rohde (2006, Definition 27). According to the definition, an allocation path is efficient if, at any time t, there exists no other feasible allocation path which improves upon the initial one with respect to the preferences of all the current and future selves following period t. In the first main result of this chapter we show that any competitive equilibrium allocation is efficient in this sense. Therefore, we present a version of the First Fundamental Welfare Theorem for economies with time-variant preferences.

Our thesis concentrates on economies with a complete market structure. Each period the current selves may freely exchange their goods and rights to future consumption on the spot and futures markets respectively, as long as their budget constraints are satisfied. Moreover, given that the agents are sophisticated, we expect that, in equilibrium, the chosen consumption streams are time-consistent, i.e., they coincide with a subgame perfect Nash equilibrium path of play of the game between different selves of every agent. In particular, our market structure allows the agents to transfer their wealth across periods strategically and enables a sophisticated interaction between the subsequent incarnations of the consumers. Our class of economies differs from those specified by Luttmer and Mariotti (2003, 2007) or Herings and Rohde (2008). However, as we argue in the main body of this chapter, this does not affect the generality of our result, since all the above concepts are allocationally equivalent. This allows us to apply our main efficiency result to various specifications of competitive equilibrium.

Our result is related to the one obtained by Herings and Rohde (2006, Theorem 30). In this paper the authors discuss a class of exchange economies with *incomplete* markets and show that, in their framework, any equilibrium allocation path is time-consistently overall Pareto efficient. In this specification of the market, the agents are not allowed to transfer their wealth across time; rather, each period t, every agent is restricted to consuming only those bundles whose value does not exceed the value of his initial endowment of goods in that period. This removes the agent's ability to save or borrow, and rules out an important channel of strategic interaction between different selves of the agent, which could affect the welfare properties of the equilibrium outcomes.

This chapter also refers to a different strand of literature which focuses on conditions under which equilibrium allocations are efficient with respect to the preferences of consumers in the initial period *only*. We will say that such allocations are *date-1 Pareto efficient*.[2] For example, Laibson (1997) has shown that once agents have access to illiquid financial instruments, they are able to commit their future incarnations to a plan which is optimal with respect to the individual preferences of the initial selves. Moreover, once we allow the agents to trade in this framework, the resulting equilibrium allocations are efficient in the above sense.

Interestingly, in some special cases, even when individual agents are unable to commit, equilibrium allocations can be date-1 Pareto efficient. This property was first observed by Barro (1999) for production economies with consumers endowed with time-separable, logarithmic preferences, and hyperbolic discounting. The result is surprising, since it implies that even though the initial incarnations cannot commit to their optimal consumption plans, there exists no other feasible allocation which could strictly improve upon the equilibrium outcome with respect to their preferences. Unfortunately, such equilibria are non-generic. As shown by Luttmer and Mariotti (2007, Proposition 3), once preferences are not homothetic, the set of equilibria and the set of allocations that are date-1 Pareto efficient intersect only at isolated points.[3] Their negative result indicates, that this form of efficiency is rare in the discussed class of models. On the other hand, our analysis presents a general welfare property of equilibrium allocations in economies with time-dependent preferences. Therefore, our positive result completes the characterisation of equilibrium outcomes.

It is worth pointing out that time-consistent overall Pareto efficiency is a weaker concept than the so called *renegotiation proofness* introduced by Luttmer and Mariotti

[2] Our use of this term follows Luttmer and Mariotti (2007, Definition 1(i)). On the other hand, Herings and Rohde (2006, Definition 10) simply call such allocations *Pareto efficient*. It is worth pointing out that date-1 Pareto efficiency and time-consistent overall Pareto efficiency are not comparable, since the latter notion is generally inefficient with respect to the initial selves *only*.

[3] In fact, Luttmer and Mariotti (2007) show that the statement is true whenever the preferences of agents are nowhere locally homothetic.

(2007, Definition 1(ii)). Using our terminology, an allocation path is renegotiation proof if it is time-consistently overall Pareto efficient, and there exists no other time-consistently overall Pareto efficient allocation path which dominates the initial one with respect to preferences of the initial selves *only*. However, as we show in Section 1.3.2, the opposite implication does not hold, as the concept of renegotiation proofness is more demanding than time-consistent overall Pareto efficiency.

What is interesting, is that in some cases the three notions of efficiency, i.e., time-consistent overall Pareto efficiency, date-1 Pareto efficiency, and renegotiation proofness, coincide. We discuss one such prominent example in Section 1.5.4.

In the second part of the chapter we present conditions under which every time-consistently overall Pareto efficient allocation can be represented by a solution to a social welfare maximisation problem. The result refers to the characterization of competitive equilibria presented by Negishi (1960), who has shown that every equilibrium can be represented as a solution to a weighted social welfare maximisation problem. We extend this idea to exchange economies with time-dependent preferences, and introduce a notion of *recursive social welfare* obtained via a multi-stage optimisation problem. At each stage the social planner maximises a weighted social welfare function of the current selves, subject to him choosing amongst allocations that solve an analogous social welfare problem in all the subsequent periods. Therefore, since the choices of the social planner are determined in a similar manner to the one of the individual agents, the economy admits a form of a sophisticated representative consumer with time-dependent preferences.

The approach presented by Negishi has found a wide application to welfare economics, general equilibrium, as well as macroeconomics. For this reason, we believe that extending the idea to economies with time-dependent preferences will be useful for more applied studies of this class of economies.

The remainder of this chapter is organised as follows. In Section 1.2 we introduce our framework and the necessary notation. Then, in Section 1.3, we present the notion of time-consistent overall Pareto efficiency and compare it to the alternative concepts discussed in the literature. We state our main efficiency result in Section 1.4. Finally, Section 1.5 concerns representation of time-consistently overall Pareto efficient allocations by a solution to the recursive social welfare maximisation problem. The auxiliary results are introduced in Appendix A at the end of the thesis.

1.2 Economy with time-dependent preferences

Consider a dynamic exchange economy with a finite time-horizon T. With a slight abuse of the notation, by T we shall also denote the set of time indices, i.e., we let $T := \{1, 2, \ldots, T\}$. Moreover, we assume that the economy consists of a finite number of consumers $i \in I$.

Let $X_t = \mathbb{R}^{n_t}_+$, where $t \in T$, be a positive orthant of a n_t-dimensional Euclidean space. We shall refer to X_t as to the period t *commodity space*. Hence, n_t is the number of consumable goods at time t. Denote an arbitrary bundle of period t commodities of consumer $i \in I$ by $x^i_t \in X_t$.

Due to the dynamic nature of the economy, we find it convenient to consider paths of consumption. Let $\hat{X}_t := \times^T_{s=t} X_s$ be a *set of consumption paths* from time t to the final period T. Therefore, an element $\hat{x}^i_t \in \hat{X}_t$ is a sequence of bundles $\hat{x}^i_t := (x^i_s)^T_{s=t}$, where $x^i_s \in X_s$, for all s. We shall refer to $\hat{x}^i_t \in \hat{X}_t$ as to a consumption path of consumer i from period t onwards. In particular, $\hat{x}^i_1$ denotes a complete consumption path of consumer i from the initial time 1 till the final period T. Moreover, by definition, for any $t \in T$ and $s > t$, we have $\hat{x}^i_t = (x^i_t, \hat{x}^i_{t+1}) = (x^i_t, \ldots, x^i_{s-1}, \hat{x}^i_s)$.

Following Strotz (1955), we characterise every agent $i \in I$ by a sequence of preference relations $\{\succeq^i_t\}_{t \in T}$. We shall refer to $\succeq^i_t$ defined over $\hat{X}_t$ as to a preference relation of period t self of agent i.[4] For any $s > t$, we allow for preference relations $\succeq^i_t$ and $\succeq^i_s$ to differ over $\hat{X}_s$. That is, we admit the case in which for some $\hat{x}^i_s$, $\hat{y}^i_s \in \hat{X}_s$ and $\hat{x}^i_t = (x^i_t, \ldots, x^i_{s-1}, \hat{x}^i_s) \in \hat{X}_t$, we have $(x^i_t, \ldots, x^i_{s-1}, \hat{x}^i_s) \succeq^i_t (x^i_t, \ldots, x^i_{s-1}, \hat{y}^i_s)$ and $\hat{y}^i_s \succ^i_s \hat{x}^i_s$, which would suggest a preference reversal. In fact, the change of preferences between periods is the source of time-inconsistency in our analysis. Moreover, we assume that preferences of period t selves are not directly affected nor depend on the consumption in the preceding periods.[5] Finally, we denote the anti-symmetric and symmetric elements of $\succeq^i_t$ in the standard fashion by $\succ^i_t$ and $\sim^i_t$, for all $t \in T$. In order to make our presentation more transparent, we discuss the following example.

[4]Formally, we say that $\succeq^i_t \subset \hat{X}_t \times \hat{X}_t$.

[5]This condition is equivalent to *strong independence of past consumption* introduced by Herings and Rohde (2006, Definition 24).

Example 1.1 (Quasi-hyperbolic discounting)**.** Let $v^i : \mathbb{R}^n_+ \to \mathbb{R}$ denote an instantaneous utility function of agent $i \in I$, and δ^i, γ^i in $[0, 1)$ be respectively his long-term and present-bias discount factor. Let $X_t := \mathbb{R}^n_+$, for all $t \in T$. Hence, $\hat{X}_t = \mathbb{R}^{n(T-t+1)}_+$. Utility of period t self of consumer i is evaluated by function $u^i_t : \hat{X}_t \to \mathbb{R}$:

$$u^i_t(\hat{x}^i_t) := v^i(x^i_t) + \gamma^i \sum_{s=t+1}^{T} (\delta^i)^{s-t} v^i(x^i_s).$$

Therefore, whenever we define preferences $\{\succeq^i_t\}_{t \in T}$ such that, for all $t \in T$ and any two $\hat{x}^i_t$, $\hat{y}^i_t \in \hat{X}_t$, we have

$$\hat{x}^i_t \succeq^i_t \hat{y}^i_t \quad \text{if and only if} \quad u^i_t(\hat{x}^i_t) \geq u^i_t(\hat{y}^i_t),$$

time-separable preferences with quasi-hyperbolic discounting are embedded in our framework. Note that in this case, for any $s > t$, we have

$$u^i_t(\hat{x}^i_t) \;:=\; v^i(x^i_t) + \gamma^i \sum_{k=t+1}^{s-1} (\delta^i)^{k-t} v^i(x^i_k) \;+\; \gamma^i (\delta^i)^s \left[v^i(x^i_s) + \sum_{k=s+1}^{T} (\delta^i)^{k-s} v^i(x^i_k) \right],$$

and

$$u^i_s(\hat{x}^i_s) \;:=\; \left[v^i(x^i_s) + \gamma^i \sum_{k=s+1}^{T} (\delta^i)^{k-s} v^i(x^i_k) \right],$$

which implies that, as long as the present-bias discount factor γ^i is different from 1, period t and period s selves of agent i evaluate consumption paths starting from period s differently. Hence, the preferences of the two incarnations differ over $\hat{X}_s$.

Throughout this chapter we impose the following assumption.

Assumption 1. *For all $i \in I$ and $t \in T$, preference relation $\succeq^i_t$ is*

(i) reflexive, complete, and transitive;

(ii) locally non-satiated on X_t, i.e., for all $\hat{x}^i_t := (x^i_t, \hat{x}^i_{t+1}) \in \hat{X}_t := X_t \times \hat{X}_{t+1}$ and any $\varepsilon > 0$, there exists some $y^i_t \in X_t$ such that $\|x^i_t - y^i_t\|_{X_t} < \varepsilon$ and $(y^i_t, \hat{x}^i_{t+1}) \succ^i_t \hat{x}^i_t$, where $\| \cdot \|_{X_t}$ is a norm on X_t.[6]

With a slight abuse of the notation, we denote the cardinality of the set of consumers by I. A period t *allocation* is a vector $x_t \in X^I_t$, where $x_t := (x^i_t)_{i \in I}$. In

[6]Observe that local non-satiation of $\succeq^i_t$ on X_t is a stronger condition than local non-satiation over the whole domain $\hat{X}_t$. Clearly the former implies the latter, but the opposite implication does not hold.

addition, an *allocation path* from time t onwards will be denoted by $\hat{x}_t \in \hat{X}_t^I$, where $\hat{x}_t := (\hat{x}_t^i)_{i \in I}$. Therefore, $\hat{x}_1$ is a complete allocation path from the initial period 1 till the final date T. Moreover, as in the case of consumption bundles of individual agents, for any $t \in T$ and $s > t$, we have $\hat{x}_t = (x_t, \hat{x}_{t+1}) = (x_t, \ldots, x_{s-1}, \hat{x}_s)$.

1.2.1 Market structure and equilibrium

Consider the following structure of the trade in the economy. At the beginning of every period $t \in T$, each agent inherits a vector $\hat{a}_t^i \in \hat{X}_t$ of rights to consumption bundles from that period onwards. The current selves are allowed to trade on two types of markets: *spot markets*, where they trade the current period t consumption goods, and *futures markets*, where consumers may exchange their claims to future commodities. Therefore, given the current prices, the agents may sell their bequest and spend the acquired wealth on a new bundle of consumption goods and rights to future consumption.

Throughout this chapter we assume that prices of rights to consumption goods in a given period do not depend on the time at which they are acquired. That is, for any two periods s and s' preceding time t, the prices of claims to period t consumption traded at time s and s' are equal to the spot market prices of goods at time t. Even though the assumption seems to be restrictive, we argue in the following section that it can be imposed without loss of generality. We denote prices of period t consumption goods by $p_t \in \mathbb{R}^{n_t}_{++}$. In order to make our notation compact, we shall denote $P_t := \mathbb{R}^{n_t}_{++}$. A path of prices of commodities consumable from period t onwards is denoted by $\hat{p}_t = (p_s)_{s=t}^T \in \hat{P}_t$, where $\hat{P}_t := \times_{s=t}^T P_s$. By construction, we have $\hat{p}_t = (p_t, \hat{p}_{t+1}) = (p_t, \ldots, p_{s-1}, \hat{p}_s)$, for any t and $s > t$. In particular, $\hat{p}_1 \in \hat{P}_1$ denotes a sequence of prices of all goods consumable between periods 1 and T.

We construct the optimisation problem of time t self of agent i as follows. At the beginning of the final period T, the agent inherits the rights to consumption a_T^i acquired at time $T-1$. The claims can be exchanged on the spot market for a bundle of the current consumption goods. The disposable wealth of period T self is equal to the value of the inherited rights to consumption $p_T \cdot a_T^i$, and so time T budget set is defined by values of correspondence $B_T : P_T \times X_T \rightrightarrows X_T$,

$$B_T(p_T, a_T^i) := \left\{ x_T^i \in X_T : p_T \cdot x_T^i \leq p_T \cdot a_T^i \right\}.$$

Since there is no consumption taking place beyond time T, there are no futures markets in the ultimate period. Therefore, the set of choices of period T self is equivalent to the set of the greatest elements of $B_T(p_T, a^i_T)$ with respect to $\succeq^i_T$.[7] Hence, the demand is governed by values of correspondence $V^i_T : P_T \times X_T \rightrightarrows X_T$,

$$V^i_T(p_T, a^i_T) := \left\{x^i_T \in B_T(p_T, a^i_T) : x^i_T \succeq^i_T y^i_T \;\text{ for all }\; y^i_T \in B_T(p_T, a^i_T)\right\}.$$

Given a path of prices from time $T-1$ onwards, $\hat{p}_{T-1}$, the total wealth of consumer i in the penultimate period is equal to the value of his endowment $\hat{a}^i_{T-1}$ inherited from his former self, i.e., $\hat{p}_{T-1} \cdot \hat{a}^i_{T-1}$. Hence, the budget set of agent i is determined by values of correspondence $B_{T-1} : \hat{P}_{T-1} \times \hat{X}_{T-1} \rightrightarrows \hat{X}_{T-1}$,

$$B_{T-1}(\hat{p}_{T-1}, \hat{a}^i_{T-1}) := \left\{(x^i_{T-1}, a^i_T) \in \hat{X}_{T-1} : p_{T-1} \cdot x^i_{T-1} + p_T \cdot a^i_T \leq \hat{p}_{T-1} \cdot \hat{a}^i_{T-1}\right\}.$$

In this chapter we analyse economies where sophisticated agents are endowed with time-dependent preferences and have no commitment technology. Hence, while determining their consumption paths, consumers can correctly predict preferences and choices of their future incarnations, but cannot commit to any consumption plan. This implies, that while acquiring (x^i_{T-1}, a^i_T), the agents take into account that the actual consumption taking place at time T must belong to $V^i_T(p_T, a^i_T)$. This is to say, that since period T self is not committed to any plan, he will choose the most preferable bundle from his budget set given the inherited vector of consumption rights a^i_T. Therefore, the problem of the consumer in the preceding period $T-1$ is to maximise his current preferences over the set of affordable, time-consistent consumption paths, i.e., vectors $\hat{x}^i_{T-1} = (x^i_{T-1}, x^i_T)$ in $B_{T-1}(\hat{p}_{T-1}, \hat{a}_{T-1})$ such that $x^i_T \in V^i_T(p_T, x^i_T)$. The set of all such consumption paths is determined by values of correspondence $F^i_{T-1} : \hat{P}_{T-1} \times \hat{X}_{T-1} \rightrightarrows \hat{X}_{T-1}$,

$$F^i_{T-1}(\hat{p}_{T-1}, \hat{a}^i_{T-1}) := \left\{(x^i_{T-1}, x^i_T) \in B_{T-1}(\hat{p}_{T-1}, \hat{a}^i_T) : x^i_T \in V^i_T(p_T, x^i_T)\right\}.$$

The elements of $F^i_T(\hat{p}_T, \hat{a}^i_T)$ are time-consistent in the sense, that once period $T-1$ self acquires $a^i_T = x^i_T$ rights to period T commodities, in the following period the current self has no incentive to re-trade the inherited consumption rights. Hence,

[7] For a binary relation $\succeq$ defined over some set X, we consider x to be a greatest element of X with respect to $\succeq$, or simply a $\succeq$-greatest element of X, if $x \in X$ and for all y in X, we have $x \succeq y$.

the consumption plan determined in period $T-1$ will actually be implemented at time T. This allows to define the demand correspondence $V^i_{T-1} : \hat{P}_{T-1} \times \hat{X}_{T-1} \rightrightarrows \hat{X}_{T-1}$,

$$V^i_{T-1}(\hat{p}_{T-1}, \hat{a}^i_{T-1}) := \left\{ \hat{x}^i_{T-1} \in F^i_{T-1}(\hat{p}_{T-1}, \hat{a}^i_{T-1}) : \hat{x}^i_{T-1} \succeq^i_{T-1} \hat{y}^i_{T-1} \text{ for all } \hat{y}^i_{T-1} \in F^i_{T-1}(\hat{p}_{T-1}, \hat{a}^i_{T-1}) \right\},$$

with values that determine the set of optimal, time-consistent choices of agent i.

By backward induction, it is possible to determine the set of all affordable and time-consistent consumption paths at any period $t \in T$. At the beginning of time t, every consumer is in possession of a vector of rights to consumption $\hat{a}^i_t = (a^i_s)^T_{s=t}$, where $a^i_s \in X_s$, which was inherited from the preceding period $t-1$. The budget set is then determined by values of correspondence $B_t : \hat{P}_t \times \hat{X}_t \rightrightarrows \hat{X}_t$, where

$$B_t(\hat{p}_t, \hat{a}^i_t) := \left\{ \hat{x}^i_t \in \hat{X}_t : \hat{p}_t \cdot \hat{x}^i_t \leq \hat{p}_t \cdot \hat{a}^i_t \right\}.$$

Hence, the set of affordable and time-consistent consumption paths is given by the values of correspondence $F^i_t : \hat{P}_t \times \hat{X}_t \rightrightarrows \hat{X}_t$,

$$F^i_t(\hat{p}_t, \hat{a}^i_t) := \left\{ (x^i_t, \hat{x}^i_{t+1}) \in B_t(\hat{p}_t, \hat{a}^i_t) : \hat{x}^i_{t+1} \in V^i_{t+1}(\hat{p}_{t+1}, \hat{x}^i_{t+1}) \right\},$$

where $V^i_{t+1}(\hat{p}_{t+1}, \hat{x}^i_{t+1})$ is the set of optimal, time-consistent choices of the following, period $t+1$ self. Being consistent with our recursive structure, the set is determined by values of correspondence $V^i_t : \hat{P}_t \times \hat{X}_t \rightrightarrows \hat{X}_t$,

$$V^i_t(\hat{p}_t, \hat{a}^i_t) := \left\{ \hat{x}^i_t \in F^i_t(\hat{p}_t, \hat{a}^i_t) : \hat{x}^i_t \succeq^i_t \hat{y}^i_t \text{ for all } \hat{y}^i_t \in F^i_t(\hat{p}_t, \hat{a}^i_t) \right\}.$$

Correspondences F^i_t and V^i_t are constructed in the following way. Given time $t \in T$, a path of prices $\hat{p}_t$, and a sequence of rights to consumption from time t onwards, $\hat{a}^i_t$, we determine the set of all affordable consumption paths from period t on, denoted by $B_t(\hat{p}_t, \hat{a}^i_t)$. Every element $\hat{x}^i_t = (x^i_t, \hat{x}^i_{t+1})$ of the set consists of the current consumption bundle x^i_t and a sequence of consumption rights/bundles following period t, $\hat{x}^i_{t+1}$. In order to make sure that $\hat{x}^i_t \in F^i_t(\hat{p}_t, \hat{a}^i_t)$, we need to guarantee that $\hat{x}^i_{t+1}$ is a solution to the optimisation problem of the subsequent incarnation of agent i, given that he inherits $\hat{x}^i_{t+1}$. Only then the future self has no incentive to choose a different consumption path when time $t+1$ arrives. In other words, as long as $\hat{x}^i_{t+1}$ belongs to $V^i_{t+1}(\hat{p}_{t+1}, \hat{x}^i_{t+1})$, period $t+1$ agent cannot strictly benefit from re-trading $\hat{x}^i_{t+1}$.

Finally, set $V_t^i(\hat{p}_t, \hat{a}_t^i)$ consists of $\succeq_t^i$-greatest elements of $F_t^i(\hat{p}_t, \hat{a}_t^i)$. Hence, it contains the most preferable, affordable, and time-consistent consumption bundles from the perspective of period t self.[8]

We assume that in the initial period $t = 1$ the vector of rights to consumption $\hat{a}_1^i$ from that period onwards is equivalent to the initial endowment $(e_t^i)_{t \in T}$ of agent i, where $e_t^i \in X_t^i$ is a claim to commodities consumable at time t. Hence, $\hat{a}_1^i \equiv (e_t^i)_{t \in T}$.

Definition 1. *A* complete competitive equilibrium *is a pair of an allocation path $\hat{x}_1^*$ and prices $\hat{p}_1^*$ such that*

(i) every self of every agent chooses an optimal, time-consistent consumption plan, i.e., for all $t \in T$ and $i \in I$, we have $\hat{x}_t^{*i} \in V_t^i(\hat{p}_t^*, \hat{x}_t^{*i})$;

(ii) markets clear, i.e., for all $t \in T$, we have $\sum_{i \in I} x_t^{*i} = \sum_{i \in I} e_t^i$.

In the following section we discuss in detail our definition of equilibrium, and compare it to other concepts presented in the literature.

1.2.2 Comments on the notion of equilibrium

Our analysis concentrates on a complete market structure similar to the one discussed in Herings and Rohde (2008, Definition 5.2). Each period t, the spot and futures markets are opened, which allows the current selves to trade all the commodities available at time t or at any future date. In particular, this enables the agents to revise their consumption plans according to their current preferences. Since the agents are sophisticated and take into account the choices of their future incarnations, we require that in equilibrium the consumers formulate time-consistent plans which are carried out by the following selves. Moreover, we assume that the agents can correctly predict both the prices quoted on all the future commodity markets, as well as the endowments that will be delivered at any date $t \in T$.

There are two substantial variations which make our definition different from the equilibrium specification of Herings and Rohde (2008). First of all, we assume that prices p_t of claims to period t consumption are independent of the time at which they are traded. In other words, for any two periods s and s' preceding time t,

[8] This specification is equivalent to the one introduced by Strotz (1955). However, it is a different formulation from the one investigated by Harris and Laibson (2001), or more recently by Balbus, Reffett, and Woźny (2014) in the infinite dimensional framework.

prices of claims to period t consumption traded at time s and s' are equal to p_t. Hence, we assume that the ratios of prices on all the markets remain unchanged as the time progresses. Moreover, we require that all the prices can be expressed with respect to a single numeraire, i.e., a single good in one specific period. Therefore, the assumption suggests that the prices are determined only once at the initial date and remain unchanged till the end of time. This makes our concept of general equilibrium closely related to the notion of Arrow-Debreu competitive equilibrium.

Second of all, we require that in equilibrium, at any time $t \in T$, the vector of rights to consumption inherited by the current self from the preceding period, is equal to the actual consumption path $\hat{x}_t^*$. Clearly, this requirement implies that there is no trade taking place beyond the initial period, as none of the agents has an incentive to alter the inherited claims. Hence, even though the markets reopen every period, there is no exchange between agents past date $t = 1$.

Surprisingly, even though the two assumptions seem to be restrictive, whenever Assumption 1 is satisfied,[9] we may impose them without any loss of generality. This is to say, that our definition of complete competitive equilibrium presented in Definition 1 is allocationally equivalent to the sophisticated complete equilibrium introduced by Herings and Rohde (2008, Definition 5.2), in which no conditions are imposed on the variation of prices and portfolios of assets. This observation is implied by Theorem 5.3 of the same paper, in which the authors state that even if we allow for the spot and futures market prices of commodities to vary in an unrestricted manner, in a sophisticated complete equilibrium their ratios have to be equal. Therefore, the equilibrium requires that the future market prices to consumption at time t, quoted at any time s, are equal to the time t spot market prices (up to a constant). The result is driven by the arbitrage opportunities that are created whenever the above condition is violated. In particular, this implies that we can always normalise the prices with respect to one common numeraire, without affecting the equilibrium allocation. This makes our assumption superfluous as long as our interest is concentrated on the final allocation of consumption goods.

Given the above observation, our second assumption can also be imposed without affecting the set of equilibrium allocations. Clearly, as the equilibrium spot and fu-

[9]Note that under Assumption 1 our framework satisfies the definition of a *locally non-satiated economy* analysed by Herings and Rohde (2008).

tures market prices remain unchanged as the time progresses, the value of the assets inherited at time t has to be equal to the path of consumption goods from that period onwards, for any $t \in T$. Otherwise, under local non-satiation of preferences, there would exists at least one self of a consumer who could benefit from altering the equilibrium consumption plan. However, this would violate the definition of equilibrium. Therefore, even though our specification of complete competitive equilibrium seems to be restrictive, the outcomes in our framework coincide with the ones implied by the setting introduced by Herings and Rohde (2008).

Interestingly, our main result can be directly extended to models with the sequentially complete market structure, in which the consumers are only allowed to trade on the current spot market and one-period ahead contingent commodities. Therefore, the poorer asset structure does not affect the welfare properties of equilibrium allocations. This is true due to Herings and Rohde (2008, Theorem 5.10), who show that the notion of sophisticated complete equilibrium discussed above is allocationally equivalent to the sophisticated *sequentially* complete equilibrium (see Herings and Rohde, 2008, Definition 4.2). Therefore, by the argument presented above, this notion also has to coincide with the one presented in Definition 1. Since our main result refers only to the final distribution of consumption goods, we need not be concerned with the process leading to it.

Regarding the framework of Herings and Rohde (2006, Definition 11), our economy differs in one important aspect. Herings and Rohde characterise the optimisation problem of consumers in a way which does not allow them to spend on the current period t consumption more than the value of their initial endowment of period t goods. In other words, in every period t, the budget set of agent i consists only of these consumption paths $\hat{x}^i_t = (x^i_s)^T_{s=t}$ for which $p_s \cdot x^i_s \leq p_s \cdot e^i_s$, for all s. This condition rules out the possibility of transferring wealth across periods strategically, which is the essence of the model discussed in our approach. Since the structure of the economy differs substantially, the equilibrium allocations need not coincide.

Finally, we focus solely on the welfare properties of equilibrium allocations and do not discuss existence issues. Nevertheless, sufficient conditions for equilibrium existence can be found in Luttmer and Mariotti (2003, 2006).

1.3 Efficiency and time-dependent preferences

In the following section we reintroduce the definition of time-consistent overall Pareto efficiency proposed by Herings and Rohde (2008, Definition 27). Then, we compare this concept to other notions of efficiency discussed in the literature concerning economies with time-dependent preferences.

1.3.1 Time-consistent overall Pareto efficiency

We say that period t allocation $x_t := (x_t^i)_{i \in I} \in X_t^I$ is *feasible* if $\sum_{i \in I} x_t^i \leq \sum_{i \in I} e_t^i$. An allocation path $\hat{x}_t := (x_s)_{s=t}^T \in \hat{X}_t^I$ is feasible if, for every $s \geq t$, allocation x_s is feasible. Denote the set of feasible allocation paths $\hat{x}_t$ by E_t. First, we define the notion of *post-t efficiency*, which shall become useful in the remainder of this section.

Definition 2 (Post-t efficiency)**.** *A feasible allocation path $\hat{x}_1$ is* post-t efficient*, for some $t \in T$, if there exists no other feasible allocation path $\hat{y}_1$ such that, for all $i \in I$ and $s \geq t$, we have $\hat{y}_s^i \succeq_s^i \hat{x}_s^i$, and $\hat{y}_s^i \succ_s^i \hat{x}_s^i$, for some $i \in I$ and some $s \geq t$.*

A feasible allocation path $\hat{x}_1$ is post-t efficient if there exists no other feasible allocation path which can improve upon $\hat{x}_1$ with respect to the preferences of all the agents and their different selves from period t onwards. Building up on the previous definition, we present the notion of efficiency which is central to this part of the thesis.

Definition 3 (Time-consistent overall Pareto efficiency)**.** *A feasible allocation path is* time-consistently overall Pareto efficient *if it is post-t efficient for all $t \in T$.*

The main idea behind the above definition of efficiency refers to a form of time-consistency of optimal allocations. Consider a path $\hat{x}_1$ which is post-1 efficient but *not* time-consistently overall Pareto efficient. This implies that there is no other feasible allocation path which makes all the selves of all agents weakly better off, and at least some of them (i.e., at least one self of one agent) strictly better off.

Assume that period 1 has passed and consumers find themselves in period 2. Since period 1 selves are no longer present in the economy, the remaining selves following date 1 might be willing to alter the allocation of consumption in the remaining periods. As period 1 preferences are no longer taken into consideration, there might exist a distribution of goods which improves upon the previously determined allocation from the perspective of the remaining selves. The idea of time-consistent overall Pareto efficiency is to exclude such cases by imposing an additional condition, according to which the path of allocations is also post-2 efficient. Therefore, there exists no other feasible allocation path which can improve upon the original allocation with respect to preferences of the remaining selves: $\{\succeq^i_s\}_{i\in I, s\geq 2}$. This way, even though period 1 selves are no longer present in the economy, the remaining incarnations are not willing to change the previously determined distribution of goods.

Clearly, once time $t = 2$ has passed, the remaining selves from period 3 onwards face the very same problem of time-consistency of the efficient allocation. Once again, the above notion of efficiency excludes the issue, by requiring that the allocation is jointly post-3 efficient, etc. This way, regardless of the progress of time, the remaining selves can never improve upon the time-consistently overall Pareto efficient allocation.

One can also interpret time-consistent overall Pareto efficiency from the point of view of the final period. Assume that allocation path $\hat{x}_1 = (x_1, x_2, \ldots, x_T)$ is efficient in the above sense. In particular, this means that it is post-T efficient. Hence, x_T is Pareto efficient with respect to preferences $\{\succeq^i_T\}_{i\in I}$. Let

$$R_T(x_T) := \left\{y_T \in X^I_T : y_T \text{ is feasible and } y^i_T \sim^i_T x^i_T, \text{ for all } i \in I\right\}$$

be the set of all period T feasible allocations which are Pareto equivalent to x_T with respect to $\{\succeq^i_T\}_{i\in I}$. Recall that time-consistent overall Pareto efficiency implies that $\hat{x}_1$ is also post-$(T-1)$ efficient. Hence, there exists no other feasible $\hat{y}_{T-1} = (y_{T-1}, y_T)$ such that $y_T \in R_T(x_T)$ and $\hat{y}^i_{T-1} \succeq^i_{T-1} \hat{x}^i_{T-1}$, for all $i \in I$, while $\hat{y}^i_{T-1} \succ^i_{T-1} \hat{x}^i_{T-1}$, for some i. Otherwise, it would be possible to choose a consumption path which is Pareto equivalent to the original allocation for date T selves, but is Pareto preferred by period $T-1$ incarnations.

Using the recursive structure of the efficient allocation, we know that an analogue property is satisfied in each period. Hence, given

$$R_t(\hat{x}_t) := \left\{\hat{y}_t \in \hat{X}^I_t : \hat{y}_t \text{ is feasible and } \hat{y}^i_s \sim^i_s \hat{x}^i_s, \text{ for all } i \in I \text{ and } s \geq t\right\},$$

while determining whether the allocation path $\hat{x}_1 = (x_1, x_2 \ldots, \hat{x}_t)$ is time-consistently overall Pareto efficient, we need to make sure that, each period t, there exists no other feasible sub-path $\hat{y}_t = (y_t, \hat{y}_{t+1})$ such that $\hat{y}_{t+1} \in R_{t+1}(\hat{x}_{t+1})$, and $\hat{y}_t$ is strictly Pareto preferred to $\hat{x}_t$ by period t selves. Once again, this captures our understanding of time-consistency of efficient allocations.

1.3.2 Alternative notions of efficiency

The existing discussion on welfare properties of economies with time-dependent preferences concentrates around two main concepts of efficiency, namely date-t Pareto efficiency (in particular, date-1 Pareto efficiency) and renegotiation proofness. In general, the two notions do not coincide with the one specified in Definition 3. In the following section we discuss in detail the main differences between the three concepts. First, we formally reintroduce the two alternative notions of efficiency.[10]

Definition 4. *(i) A feasible allocation path $\hat{x}_1$ is* date-t Pareto efficient*, for some $t \in T$, if there exists no other feasible allocation path $\hat{y}_1$ such that $\hat{y}^i_t \succeq^i_t \hat{x}^i_t$, for all $i \in I$, and $\hat{y}^i_t \succ^i_t \hat{x}^i_t$ for some i.*[11]

(ii) Let R be the set of all feasible, time-consistently overall Pareto efficient allocation paths. Allocation path $\hat{x}_1 \in R$ is renegotiation proof *if there exists no other $\hat{y}_1 \in R$ such that $\hat{y}^i_1 \succeq^i_1 \hat{x}^i_1$, for all $i \in I$, and $\hat{y}^i_1 \succ^i_1 \hat{x}^i_1$ for some i.*

The concept of date-t Pareto efficiency is not comparable to time-consistent overall Pareto efficiency. First of all, since the latter notion takes into account the preferences of all the different selves of all agents, in general it is possible to improve the welfare of period t incarnations once the remaining selves are not taken into consideration. Hence, time-consistent overall Pareto efficiency does not imply date-t Pareto efficiency.

To see that the opposite implication also fails to hold, suppose that allocation path $\hat{x}_1$ is date-t Pareto efficient. There are two possible ways in which the allocation path may violate the definition of time-consistent overall Pareto efficiency. First of all, there might exist some other feasible allocation $\hat{y}_1$ which is Pareto equivalent to $\hat{x}_1$ with respect to period t selves, i.e., $\hat{y}^i_t \sim^i_t \hat{x}^i_t$, for all $i \in I$, but strictly improves the welfare of the remaining incarnations, i.e., $\hat{y}^i_s \succeq^i_s \hat{x}^i_s$, for all $i \in I$ and $s \neq t$, and

[10] Renegotiation proofness was defined by Luttmer and Mariotti (2007) only for two period economies. Therefore, we extend the definition in a way we think is accordant with the intuition of the authors.

[11] Note that date-t efficiency and post-t efficiency are different concepts. The former notion takes into account the preferences of period t selves *only*, while the latter considers the welfare of all the selves from period t onwards.

$\hat{y}^i_s \succ^i_s \hat{x}^i_s$, for some i and $s \neq t$. Second of all, even if the above Pareto improvement is not possible, the allocation path might fail to be post-s efficient for some $s > t$. In other words, there might exist some feasible allocation path $\hat{y}_1$ and some time period $s > t$ such that $\hat{y}^i_r \succeq^i_r \hat{x}^i_r$, for all $r \geq s$ and $i \in I$, while $\hat{y}^i_r \succ^i_r \hat{x}^i_r$, for some $r \geq s$ and i. Once the allocation fails to be post-s efficient for some $s > t$, it violates the definition of time-consistent overall Pareto efficiency.

Clearly, time-consistent overall Pareto efficiency is a weaker concept than renegotiation proofness. Consider the two-period case as in Luttmer and Mariotti (2007), with $T = \{1, 2\}$. Denote the set of all period 2 Pareto efficient allocations by

$$R_2 := \{x_2 \in X_2^I : x_2 \text{ is feasible and there is no feasible } y_2 \in X_2^I \\ \text{such that } y^i_2 \succeq^i_2 x^i_2, \text{ for all } i \in I, \text{ and } y^i_2 \succ^i_2 x^i_2, \text{ for some } i\}.$$

Given the definition by Luttmer and Mariotti (2007, Definition 1(ii)), a feasible allocation path $\hat{x}_1 := (x_1, x_2)$ is renegotiation proof if: (a) $x_2 \in R_2$, and (b) there exist no other feasible allocation path $\hat{y}_1 := (y_1, y_2)$, with $y_2 \in R_2$, such that $\hat{y}^i_1 \succeq^i_1 \hat{x}^i_1$, for all $i \in I$, and $\hat{y}^i_1 \succ^i_1 \hat{x}^i_1$, for some i. Clearly, since x_2 belongs to R_2, allocation $\hat{x}^i_1$ is post-2 efficient. In addition, there is no other allocation path $\hat{y}_1 := (y_1, y_2)$, with $y_2 \in R_2$, hence, no other post-2 efficient allocation path, which could Pareto improve upon $\hat{x}_1$ with respect to period 1 preferences. Therefore, any improvement in the welfare of the initial selves would have to make worst off at least some incarnations in the final period, which makes $\hat{x}_1$ a time-consistently overall Pareto efficient allocation.

To see why the opposite implication does not hold, suppose that $\hat{x}_1 = (x_1, x_2)$ is time-consistently overall Pareto efficient. This implies that (a') the allocation path is post-2 efficient, hence, $x_2 \in R_2$, and (b') $\hat{x}_1$ is post-1 efficient. Let

$$R_2(x_2) := \left\{y_2 \in X_2^I : y_2 \text{ is feasible and } y^i_2 \sim^i_2 x^i_2, \text{ for all } i \in I\right\}$$

be the set of allocations that are Pareto equivalent to x_2 with respect to period 2 selves. Clearly, we have $R_2(x_2) \subseteq R_2$. Condition (b') implies that there is no other feasible allocation path $\hat{y}_1 = (y_1, y_2)$ such that $y_2 \in R_2(x_2)$ and $\hat{y}^i_1 \succeq^i_1 \hat{x}^i_1$, for all $i \in I$, while $\hat{y}^i_1 \succ^i_1 \hat{x}^i_1$. Clearly, (a) and (a') are equivalent. However, since $R_2(x_2) \subset R_2$, requirement (b') is weaker than (b).

1.4 Efficiency of competitive equilibria

In the following section we present and prove the first main result of this chapter concerning the efficiency of equilibrium allocations. We begin by stating the theorem.

Theorem 1.1. *Under Assumption 1, any complete competitive equilibrium allocation path is time-consistently overall Pareto efficient.*

The above theorem states that, in particular, there exists no other feasible allocation path which can improve upon the equilibrium outcome with respect to the preferences of all the different selves of every consumer. However, since the outcome is time-consistently overall Pareto efficient, it additionally implies that as the time progresses and the initial selves are successively excluded from the economy, there exists no allocation which could improve the welfare of the remaining incarnations. Therefore, equilibrium outcomes do not give much room for improvement. Clearly, an equilibrium allocation is usually neither date-1 Pareto efficient, nor renegotiation proof. Nevertheless, any change in the allocation of goods would make worst off at least some of the incarnations.

In order to prove Theorem 1.1, we need to show that any equilibrium allocation path is post-t efficient for all $t \in T$. To achieve his goal, we apply an inductive argument. Since the proof is rather extensive, we find it convenient to present it via several lemmas. In the first result we show that any equilibrium allocation path is post-T efficient, which at the same time will constitute the base step for our argument. Throughout this section we consider Assumption 1 to be satisfied. Moreover, we assume that pair $(\hat{x}_1^*, \hat{p}_1^*)$ constitutes a complete competitive equilibrium.

Lemma 1.1. *Allocation path $\hat{x}_1^*$ is post-T efficient and for any feasible allocation path $\hat{y}_1 \in \hat{X}_1$ such that $y_T^i \sim_T^i x_T^{*i}$, for all $i \in I$, we have $p_T \cdot y_T^i = p_T \cdot x_T^{*i}$, for all $i \in I$.*

Proof. First, we show that for any $i \in I$ and $y_T^i \in X_T$, we have that (i) $y_T^i \succeq_T^i x_T^{*i}$ implies $p_T^* \cdot y_T^i \geq p_T^* \cdot x_T^{*i}$, while (ii) $y_T^i \succ_T^i x_T^{*i}$ implies $p_T^* \cdot y_T^i > p_T^* \cdot x_T^{*i}$. We prove both claims by contradiction. Suppose that $y_T^i \succeq_T^i x_T^{*i}$ and $p_T^* \cdot y_T^i < p_T^* \cdot x_T^{*i}$. By Assumption 1, there is some $z_T^i \in B_T(p_T^*, x_T^{*i})$ such that $z_T^i \succ_T^i y_T^i \succeq_T^i x_T^{*i}$. This contradicts that $x_T^{*i} \in V_T^i(p_T^*, x_T^{*i})$. To show that (ii) holds, assume $y_T^i \succ_T^i x_T^{*i}$. By claim (i), we know that $p_T^* \cdot y_T^i \geq p_T^* \cdot x_T^{*i}$. Whenever the condition holds with equality, we have $y_T^i \in B_T(p_T^*, x_T^{*i})$, which contradicts that $x_T^{*i} \in V_T^i(p_T^*, x_T^{*i})$.

In order to prove the first part of the lemma, suppose that $\hat{x}_1$ is not post-T efficient. Hence, there is some feasible allocation path $\hat{y}_1$ such that $y_T^i \succeq_T^i x_T^{*i}$, for all $i \in I$, and $y_T^i \succ_T^i x_T^{*i}$ for some i. Claims (i) and (ii) imply that, for all $i \in I$, we have $p_T^* \cdot y_T^i \geq p_T^* \cdot x_T^{*i}$, while for some i the inequality is strict. Therefore,

$$p_T^* \cdot \sum_{i \in I} e_T^i \geq p_T^* \cdot \sum_{i \in I} y_T^i > p_T^* \cdot \sum_{i \in I} x_T^{*i} = p_T^* \cdot \sum_{i \in I} e_T^i,$$

where the weak inequality follows from feasibility of y_T, while the equality is implied by the market clearing condition. Clearly, we reach a contradiction.

Finally, to prove the second part of the lemma, take any feasible allocation path $\hat{y}_1$ such that $y_T^i \sim_T^i x_T^{*i}$, for all $i \in I$. By claim (i), this implies that $p_T^* \cdot y_T^i \geq p_T^* \cdot x_T^{*i}$, for all $i \in I$. Whenever there exists some $i \in I$ for which the above inequality is strict, we obtain a contradiction analogous to the one above. The proof is complete. □

Lemma 1.1 is a simple reformulation of the First Fundamental Welfare Theorem. Since in the final period our model is a static Arrow-Debreu economy, any equilibrium allocation is Pareto efficient, hence, post-T efficient. However, the result highlights one important property of an equilibrium. Namely, for any feasible allocation which is Pareto equivalent to the equilibrium outcome, the value of individual bundles, given the equilibrium prices, is equal to those chosen in the equilibrium. In fact, as we show in the remainder of this section, a similar property is satisfied in our dynamic framework with time-dependent preferences. We proceed with the following claim.

Lemma 1.2. *For any $t \in T$, $i \in I$, take some $\hat{y}_t^i \in \hat{X}_t$ such that for all $s \geq t+1$, we have $\hat{y}_s^i \sim_s^i \hat{x}_s^{*i}$ and $p_s^* \cdot y_s^i = p_s^* \cdot x_s^{*i}$. Then, $\hat{y}_t^i \succeq_t^i \hat{x}_t^{*i}$ implies $p_t^* \cdot y_t^i \geq p_t^* \cdot x_t^{*i}$, while $\hat{y}_t^i \succ_t^i \hat{x}_t^{*i}$ implies $p_t^* \cdot y_t^i > p_t^* \cdot x_t^{*i}$.*

Proof. We prove the first part of the claim by contradiction. Assume that $\hat{y}_t^i \succeq_t^i \hat{x}_t^{*i}$, and $p_t^* \cdot y_t^i < p_t^* \cdot x_t^{*i}$. By assumption, for all $s \geq t+1$, we have $p_s^* \cdot y_s^i = p_s^* \cdot x_s^{*i}$. In particular, this is equivalent to $\hat{p}_{t+1}^* \cdot \hat{y}_{t+1}^i = \hat{p}_{t+1}^* \cdot \hat{x}_{t+1}^{*i}$. Therefore, given the initial inequality, it must be that $p_t^* \cdot y_t^i + \hat{p}_{t+1}^* \cdot \hat{y}_{t+1}^i < p_t^* \cdot x_t^{*i} + \hat{p}_{t+1}^* \cdot \hat{x}_{t+1}^{*i}$. Hence, $\hat{y}_t^i \in B_t(\hat{p}_t^*, \hat{x}_t^{*i})$. Assumption 1(ii) implies that there exists some $z_t^i \in X_t$ such that

$(z_t^i, \hat{y}_t^i) \in B_t(\hat{p}_t^*, \hat{x}_t^{*i})$ and $(z_t^i, \hat{y}_{t+1}^i) \succ_t^i \hat{y}_t^i$, which by transitivity of $\succeq_t^i$ implies that $(z_t^i, \hat{y}_{t+1} i) \succ_t^i \hat{x}_t^{*i}$.

By Lemma A.1 (see Appendix A), we have $V_{t+1}^i(\hat{p}_{t+1}^*, \hat{x}_{t+1}^{*i}) = V_{t+1}^i(\hat{p}_{t+1}^*, \hat{y}_{t+1}^i)$. In addition, Lemma A.2 (also in Appendix A) implies that $\hat{y}_{t+1}^i \in V_{t+1}^i(\hat{p}_{t+1}^*, \hat{x}_{t+1}^{*i})$. Given the previous observation, we conclude that $\hat{y}_{t+1}^i \in V_{t+1}^i(\hat{p}_{t+1}^*, \hat{y}_{t+1}^i)$. Therefore, consumption path $(z_t^i, \hat{y}_{t+1}^i)$ belongs to $F_t^i(\hat{p}_t^*, \hat{x}_t^{*i})$. However, as $(z_t^i, \hat{y}_{t+1}^i) \succ_t^i \hat{x}_t^{*i}$, this contradicts that $\hat{x}_t^{*i}$ belongs to $V_t^i(\hat{p}_t^*, \hat{x}_t^{*i})$. Hence, it must be that $\hat{y}_1^i \succeq_t^i \hat{x}_t^{*i}$ implies $p_t^* \cdot y_t^i \geq p_t^* \cdot x_t^{*i}$.

To prove the second part of the claim, assume that $\hat{y}_t^i \succ_t^i \hat{x}_t^{*i}$. By the argument presented in the first part of the lemma, we know that $p_t^* \cdot y_t^i \geq p_t^* \cdot x_t^{*i}$. Suppose that for some $i \in I$, we have $p_t^* \cdot y_t^i = p_t^* \cdot x_t^{*i}$. Then $\hat{y}_t^i \in B_t(\hat{p}_t^*, \hat{x}_t^{*i})$. Moreover, by Lemma A.1, we have $\hat{y}_{t+1}^i \in V_{t+1}^i(\hat{p}_{t+1}^*, \hat{x}_{t+1}^{*i})$, while Lemma A.2 implies that $\hat{y}_{t+1}^i \in V_t^i(\hat{p}_{t+1}^*, \hat{y}_{t+1}^i)$. Therefore, we have $\hat{y}_t^i \in F_t^i(\hat{p}_t^*, \hat{x}_t^{*i})$. However, as $\hat{y}_t^i \succ_t^i \hat{x}_t^{*i}$, this contradicts that $\hat{x}_t^{*i} \in V_t^i(\hat{p}_t^*, \hat{x}_t^{*i})$. Hence, we conclude that $\hat{y}_t^i \succ_t^i \hat{x}_t^{*i}$ implies $p_t^* \cdot y_t^i > p_t^* \cdot x_t^{*i}$. □

The following lemma states a sufficient condition which allows to determine that an equilibrium allocation path is post-t efficient, for some $t \in T$. The result will play an important role in the remainder of the proof.

Lemma 1.3. *Suppose that $\hat{x}_1^*$ is post-$(t+1)$ efficient. Moreover, assume that for any feasible allocation path $\hat{y}_1$ such that $\hat{y}_s^i \sim_s^i \hat{x}_s^{*i}$, for all $i \in I$ and $s \geq t+1$, we have $p_s^* \cdot y_s^i = p_s^* \cdot x_s^{*i}$, for all $i \in I$ and $s \geq t+1$. Then, $\hat{x}_1^*$ is post-t efficient.*

Proof. We prove the claim by contradiction. Suppose that $\hat{x}_1^*$ is not post-t efficient. Therefore, there exists some feasible allocation path $\hat{y}_1$ such that, for all $i \in I$ and $s \geq t$, we have $\hat{y}_s^i \succeq_s^i \hat{x}_s^{*i}$, and for some i and $s \geq t$, $\hat{y}_s^i \succ_s^i \hat{x}_s^{*i}$.

By assumption, $\hat{x}_1^*$ is post-$(t+1)$ efficient, so it must be that $\hat{y}_s^i \sim_s^i \hat{x}_s^{*i}$, for all $i \in I$ and $s \geq t+1$. This implies that $\hat{y}_t^i \succeq_t^i \hat{x}_t^{*i}$, for all $i \in I$, and $\hat{y}_t^i \succ_t^i \hat{x}_t^{*i}$, for some i. Therefore, by Lemma 1.2, we obtain

$$p_t^* \cdot \sum_{i \in I} e_t^i \geq p_t^* \cdot \sum_{i \in I} y_t^i > p_t^* \cdot \sum_{i \in I} x_t^{*i} = p_t^* \cdot \sum_{i \in I} e_t^i,$$

where the weak inequality follows from the feasibility of $\hat{y}_1$, while the equality is implied by the market clearing condition. Contradiction. □

The final lemma allows us to determine allocation paths whose spot market values are equal to the value of the equilibrium outcome. At the same time the result constitutes the final prerequisite which is necessary to construct the inductive step.

Lemma 1.4. *Take any feasible allocation path $\hat{y}_1$ such that, for all $i \in I$ and $s \geq t+1$, we have $\hat{y}_s^i \sim_s^i \hat{x}_s^i$ and $p_s^* \cdot y_s^i = p_s^* \cdot x_s^{*i}$. Whenever $\hat{y}_t^i \sim_t^i \hat{x}_t^{*i}$, for all $i \in I$, then $p_t^* \cdot y_t^i = p_t^* \cdot x_t^{*i}$, for all $i \in I$.*

Proof. Take any feasible allocation path $\hat{y}_1$ that satisfies the thesis of the lemma. By Lemma 1.2, we know that, for all $i \in I$, whenever $\hat{y}_t^i \sim_t^i \hat{x}_t^{*i}$ then $p_t^* \cdot y_t^i \geq p_t^* \cdot x_t^{*i}$. Suppose that for some i the inequality is strict. Then

$$p_t^* \cdot \sum_{i \in I} e_t^i \geq p_t^* \cdot \sum_{i \in I} y_t^i > p_t^* \cdot \sum_{i \in I} x_t^{*i} = p_t^* \cdot \sum_{i \in I} e_t^i.$$

Analogously as in the proof of Lemma 1.3, we reach a contradiction. □

Finally, given the preliminary results, we conclude this section with our argument supporting the first main theorem of this chapter.

Proof of Theorem 1.1. Suppose that a tuple $(\hat{x}_1^*, \hat{p}_1^*)$ constitutes a complete competitive equilibrium. We need to show that allocation path $\hat{x}_1^*$ is post-t efficient for any $t \in T$. We prove the result by induction. By Lemma 1.1, we know that $\hat{x}_1^*$ is post-T efficient. Moreover, for any other feasible allocation path $\hat{y}_1$ such that $y_T^i \sim_T^i x_T^{*i}$, for all $i \in I$, we have $p_T^* \cdot y_T^i = p_T^* \cdot x_T^{*i}$.

To show the inductive step, suppose that $\hat{x}_1^*$ is post-$(t+1)$ efficient and that for any feasible allocation path $\hat{y}_t$ such that $\hat{y}_s^i \sim_s^i \hat{x}_s^{*i}$, for all $i \in I$ and $s \geq t+1$, we have $p_s^* \cdot y_s^i = p_s^* \cdot x_s^{*i}$, for all $i \in I$ and $s \geq t+1$. Then, Lemma 1.3 implies that allocation $\hat{x}_1^*$ is post-t efficient. Moreover, Lemma 1.4 states that for any feasible allocation path $\hat{y}_1$ such that $\hat{y}_s^i \sim_s^i \hat{x}_s^{*i}$, for all $i \in I$ and $s \geq t$, we have $p_s^* \cdot y_s^i = \hat{p}_s^* \cdot x_s^{*i}$, for all $i \in I$ and $s \geq t$.

By induction, we conclude that whenever $\hat{x}_1^*$ is post-T efficient, then it is post-t efficient, for any $t \in T$. Therefore, the equilibrium allocation path $\hat{x}_1^*$ is time-consistently overall Pareto efficient. □

1.5 Representation of efficient allocations

In the following section we concentrate on a representation of time-consistently overall Pareto efficient allocations by solutions to a social welfare optimisation problem. Throughout this section we impose the following condition.

Assumption 2 (Utility representation)**.** *For all $i \in I$ and $t \in T$, preference relation $\succeq_t^i$ is represented by a utility function $u_t^i : \hat{X}_t \to \mathbb{R}$. That is, for any two $\hat{x}_t^i$, $\hat{y}_t^i \in \hat{X}_t$, we have $\hat{y}_t^i \succeq_t^i \hat{x}_t^i$ if and only if $u_t^i(\hat{y}_t^i) \geq u_t^i(\hat{x}_t^i)$.*[12]

In the remainder of the section we characterize our notion of social welfare. Then, we discuss when this concept coincides with time-consistent overall Pareto efficiency presented in the preceding section.

1.5.1 Recursive social welfare

We construct our notion of social welfare function using backward induction. Recall that, for all $t \in T$, E_t denotes the set of feasible allocation paths $\hat{x}_t$. First, consider the social planner's problem in the final period $t = T$. For any non-zero weights $\alpha_T := (\alpha_T^i)_{i \in I} \in \mathbb{R}_+^I$, define set

$$\Psi_T(\alpha_T) := \operatorname*{argmax}_{x_T \in E_T} \sum_{i \in I} \alpha_T^i u_T^i(x_T^i). \tag{1.1}$$

In other words, $\Psi_T(\alpha_T)$ contains all feasible period T consumption bundles which maximise the weighted social welfare function for a fixed vector of weights $\alpha_T := (\alpha_T^i)_{i \in I}$. Since the form of the above functional is rather standard, we refrain from further discussion.

Next, consider the problem in period $t = T - 1$. Denote a path of non-zero weights from period $T-1$ onwards by $\hat{\alpha}_{T-1} := (\alpha_{T-1}, \alpha_T)$, where $\alpha_t = (\alpha_t^i)_{i \in I} \in \mathbb{R}_+^I$, $t \in \{T-1, T\}$. Define

$$\Psi_{T-1}(\hat{\alpha}_{T-1}) := \operatorname*{argmax}_{\hat{x}_{T-1} \in \Gamma_{T-1}(\alpha_T)} \sum_{i \in I} \alpha_{T-1}^i u_{T-1}^i(\hat{x}_{T-1}^i),$$

where

$$\Gamma_{T-1}(\alpha_T) := \left\{(x_{T-1}, x_T) \in E_T : x_T \in \Psi_T(\alpha_T)\right\},$$

[12] Sufficient conditions for utility representation of preferences are well-known (e.g. see Mas-Colell, Whinston, and Green, 1995, Chapter 3.C).

and $\Psi_T(\alpha_T)$ is defined as in (1.1). Therefore, set $\Psi_{T-1}(\hat{\alpha}_{T-1})$ contains all allocation paths from date $T-1$ onwards which maximise period $T-1$ social welfare functional for weights α_{T-1}, given that period T allocation x_T is a solution to the social planer's optimisation problem in the final period for weights α_T. In other words, set $\Psi_{T-1}(\hat{\alpha}_{T-1})$ contains time-consistent, welfare maximising allocations, in an environment where the social planner faces a similar time-inconsistency problem as the individual consumers.

Using backward induction, one can determine the corresponding sets $\Psi_t(\hat{\alpha}_t)$ and $\Gamma_t(\hat{\alpha}_{t+1})$ for any $t \in T$, and any path of weights $\hat{\alpha}_t := (\alpha_s)_{s=t}^T$, where $\alpha_s \in \mathbb{R}_+^I$. Define set

$$\Psi_t(\hat{\alpha}_t) := \underset{\hat{x}_t \in \Gamma_t(\hat{\alpha}_{t+1})}{\operatorname{argmax}} \sum_{i \in I} \alpha_t^i u_t^i(\hat{x}_t^i), \tag{1.2}$$

where

$$\Gamma_t(\hat{\alpha}_{t+1}) := \{(x_t, \hat{x}_{t+1}) \in E_t : \hat{x}_{t+1} \in \Psi_{t+1}(\hat{\alpha}_{t+1})\},$$

and $\Psi_{t+1}(\hat{\alpha}_{t+1})$ is defined as in (1.2) for period $t+1$ selves and the corresponding subsequence of weights $\hat{\alpha}_{t+1}$.

The construction of Ψ_t and Γ_t is similar to the construction of correspondences V_t^i and F_t^i for the optimisation problems of individual agents in Section 1.2.1. Namely, for any t, take the set E_t of feasible allocation paths following time t. By definition, for any $\hat{x}_t = (x_t, \hat{x}_{t+1}) \in E_t$, x_t is an allocation of period t consumption goods and $\hat{x}_{t+1}$ is a path of allocations following period $t+1$. In order to make sure that $\hat{x}_t \in \Gamma_t(\hat{\alpha}_t)$, we need to guarantee that the subsequence $\hat{x}_{t+1}^i$ is a solution to the corresponding social welfare optimisation problem in the following period, given the path of weights $\hat{\alpha}_{t+1}$. This way, we obtain a form of time-consistency of socially optimal allocations. That is, given that the next period social planner is guided by a different social welfare function, he is not willing to change the allocation determined in the preceding period, as it could not strictly improve the current welfare given his criterion. Finally, the social planner in period t chooses an element from the set of feasible, time-consistent sequences of allocations, that maximises his current objective.

Definition 5 (Recursive social welfare)**.** *An allocation path* $\hat{x}_1^\circ := (\hat{x}_1^{\circ i})_{i \in I} \in \hat{X}_1^I$ *maximises the* recursive social welfare*, if there exists a path of non-zero weights* $\hat{\alpha}_1 := (\alpha_t)_{t \in T}$*, where* $\alpha_t \in \mathbb{R}_+^I$*, such that* $\hat{x}_1^\circ \in \Psi_1(\hat{\alpha}_1)$*.*

By definition of the recursive social welfare, we assume that whenever there exist two distinct allocation paths $\hat{x}_t$, $\hat{y}_t \in \Psi_t(\hat{\alpha}_t)$ that solve the corresponding social welfare problem at time t, then the actual choice is always determined by the social planner in the initial period. Therefore, as in the case of the individual consumer demand, or time-consistent overall Pareto efficiency, we concentrate on socially optimal allocations that are time-consistent. In other words, the recursive social welfare allocation might be interpreted as a subgame perfect Nash equilibrium path of the game between different incarnations of the social planner.

The immediate question that follows from the above definition concerns conditions under which there exists a solution to the recursive social welfare maximisation problem. It is rather straightforward to show that, whenever function u_t^i is upper semi-continuous, for all $i \in I$ and $t \in T$, the set of recursive social welfare allocations is non-empty. In fact, under this assumption, for any $t \in T$ and any weights $\hat{\alpha}_t := (\alpha_s)_{s=t}^T$, where $\alpha_s \in \mathbb{R}_+^I$, function $\sum_{i\in I} \alpha_t^i u_t^i$ is upper semi-continuous, while set $\Gamma_t(\hat{\alpha}_t)$ is non-empty and compact. Hence, every period t social welfare optimisation problem has a solution.

An allocation maximising the recursive social welfare is a solution to a multi-stage optimisation problem, where at each stage t the social planner maximises the current period weighted social welfare function, given that the path of allocations following time t is a solution to an analogue problem in each of the subsequent periods. Therefore, the recursive social welfare is closely related to time-consistent overall Pareto efficiency, as it focuses on a form of time-consistency of optimal allocations. In the next section we present conditions under which the two notions coincide.

1.5.2 Social welfare and efficiency

First, we show conditions under which every allocation that maximises the recursive social welfare is time-consistently overall Pareto efficient.

Proposition 1.1. *If $\hat{x}_1^\circ = (x_1^{\circ i})_{i\in I} \in \hat{X}_1^I$ maximises the recursive social welfare for some strictly positive path of weights $\hat{\alpha}_T := (\alpha_t)_{t\in T}$, i.e., $\alpha_t \in \mathbb{R}_{++}^I$, for all $t \in T$, then it is time-consistently overall Pareto efficient.*

Proof. Let $\hat{x}_1^\circ = (x_1^\circ, \dots, x_T^\circ)$ maximise the recursive social welfare for some real, strictly positive path of weights $\hat{\alpha}_1$. We prove the result by induction. First, we show that $\hat{x}_1^\circ$ is post-T efficient. Assume the opposite. Then, there exists some $y_T \in E_T$ such that, for all $i \in I$, $u_T^i(y_T^i) \geq u_T^i(x_T^{\circ i})$, and for some i, $u_T^i(y_T^i) > u_T^i(x_T^{\circ i})$. Since weights α_T are strictly positive, this implies $\sum_{i\in I} \alpha_T^i u_T^i(y_T^i) > \sum_{i\in I} \alpha_T^i u_T^i(x_T^{\circ i})$, which contradicts that $x_T^\circ \in \Psi_T(\alpha_T)$ as well as $\hat{x}_1^\circ \in \Psi_1(\hat{\alpha}_1)$.

Next, take any $t \in T$ and assume that, for all $s \geq t+1$, $\hat{x}_1^\circ$ is post-s efficient. We claim that $\hat{x}_1^\circ$ is post-t efficient. Assume the opposite. Therefore, there exists some $\hat{y}_t \in E_t$ such that, for all $i \in I$ and $s \geq t$, we have $u_s^i(\hat{y}_s^i) \geq u_s^i(\hat{x}_s^{\circ i})$, and for some i and some $s \geq t$, $u_s^i(\hat{y}_s^i) > u_s^i(\hat{x}_s^{\circ i})$. By assumption, for all $s \geq t+1$, $\hat{x}_s^\circ$ is post-s efficient, so it must be that, for all $i \in I$ and $s \geq t+1$, $u_s^i(\hat{y}_s^i) = u_s^i(\hat{x}_s^{\circ i})$. This implies that $\hat{y}_{t+1} \in \Psi_{t+1}(\hat{\alpha}_{t+1})$, and so $\hat{y}_t \in \Gamma_t(\hat{\alpha}_{t+1})$. Moreover, since the weights are strictly positive, we have $\sum_{i\in I} \alpha_t^i u_t^i(\hat{y}_t^i) > \sum_{i\in I} \alpha_t^i u_t^i(\hat{x}_t^{\circ i})$, which contradicts that $\hat{x}_t^\circ \in \Psi_t(\hat{\alpha}_t)$ as well as $\hat{x}_1^\circ \in \Psi_1(\hat{\alpha}_1)$. □

Proposition 1.1 implies that, in general, a set of time-consistently overall Pareto efficient allocations can be determined via a solution to the recursive social welfare maximisation problem, as long as the weights corresponding to each self of every consumer are strictly positive. In particular, knowing the conditions under which there exists a solution to some recursive social welfare problem, we conclude that the set of time-consistently overall Pareto efficient allocations is non-empty.

In order to prove the converse result to Proposition 1.1, we need to impose some convexity assumptions on the preferences.

Assumption 3 (Concavity)**.** *For all $i \in I$, $t \in T$, and $\hat{x}_{t+1}^i \in \hat{X}_{t+1}$, utility function $u_t^i(x_t^i, \hat{x}_{t+1}^i)$ is continuous and strictly concave with respect to x_t^i.*

Finally, to obtain a sharper version of our results, we refer to one additional assumption.

Assumption 4 (Monotonicity)**.** *For all $i \in I$, $t \in T$, and $\hat{x}_{t+1}^i \in \hat{X}_t$, utility function $u_t^i(x_t^i, \hat{x}_{t+1}^i)$ is strictly increasing with respect to x_t^i.*

Observe that we do not require function $u_t^i(x_t^i, \hat{x}_{t+1}^i)$ to be continuous, concave, or monotone with respect to $\hat{x}_{t+1}^i$. In fact, apart from being well defined, we do not impose any conditions on the properties of the utility functions with respect to the second argument. We proceed with the second main theorem of this chapter.

Theorem 1.2. *Let Assumptions 2, 3 be satisfied and* $\hat{x}_1 = (\hat{x}_1^i)_{i \in I}$*, where* $\hat{x}_1^i = (x_t^i)_{t \in T}$*, be a time-consistently overall Pareto efficient allocation. There exists a non-zero path of weights* $\hat{\alpha}_1 := (\alpha_t)_{t \in T}$*, where* $\alpha_t \in \mathbb{R}_+^I$*, such that* $\Psi_1(\hat{\alpha}_1) = \{\hat{x}_1\}$*. If additionally Assumption 4 is satisfied and* x_t^i *is non-zero, for all* $i \in I$ *and* $t \in T$*, then* $\hat{\alpha}_1$ *is strictly positive, that is,* $\alpha_t \in \mathbb{R}_{++}^I$*, for all* $t \in T$*.*

Proof. Assume that $\hat{x}_1 = (x_1, \ldots, x_T)$ is a time-consistently overall Pareto efficient allocation. We prove the result by induction. First, take $t = T$. Let function $u_T : X_T^I \to \mathbb{R}^I$ be defined as $u_T(x_T) := (u_T^i(x_T^i))_{i \in I}$. Denote the image of function u_T over set E_T by $U_T' := u_T(E_T)$. Since for all $i \in I$, u_T^i is continuous and E_T is compact, U_T' is compact. Let $U_T := \{u \in \mathbb{R}^I : u \leq u_T(y_T), \text{ for some } y_T \in E_T\}$.[13] By Assumption 3, U_T is convex. Moreover, we have $U_T = U_T' - \mathbb{R}_+^I$. Hence, set U_T must be closed and bounded above.

Denote $u_T^* = u(x_T)$. By definition of $\hat{x}_1$, there exists no other $y_T \in E_T$ such that, for all $i \in I$, $u_T^i(y_T^i) \geq u_T^i(x_T^i)$, and $u_T^i(y_T^i) > u^i(x_T^i)$, for some i. Hence, it must be that $u_T^* \in \partial U_T$. By the Separating Hyperplane Theorem (see e.g. Aliprantis and Border, 2006, Theorem 7.30), there exists some non-zero vector $\alpha_T \in \mathbb{R}^I$ such that, for all $u \in U_T$, $\alpha_T \cdot u_T^* \geq \alpha_T \cdot u$. Since $U_T - \mathbb{R}_+^I \subset U_T$, it must be that $\alpha_T \in \mathbb{R}_+^I$. By construction, this implies that $x_T \in \Psi_T(\alpha_T)$. Finally, by strict concavity of u_T^i and convexity of E_T, it must be that $\Psi_T(\alpha_T) = \{x_T\}$.

Next, take any $t \in T$. Assume that there exists a path of non-zero, positive vectors $\hat{\alpha}_{t+1}$ such that $\Psi_{t+1}(\hat{\alpha}_{t+1}) = \{\hat{x}_{t+1}\}$. Clearly, in this case $\Gamma_t(\hat{\alpha}_{t+1})$ is compact and convex. Let function $u_t : \hat{X}_t^I \to \mathbb{R}^I$ be defined by $u_t(\hat{x}_t) := (u_t^i(\hat{x}_t^i))_{i \in I}$, and let $U_t' := u_t(\Gamma_t(\hat{\alpha}_{t+1}))$ denote the image of function u_t over set $\Gamma_t(\hat{\alpha}_{t+1})$. Since, for all $\hat{x}_{t+1}^i \in \hat{X}_{t+1}$, function $u_t^i(x_t^i, \hat{x}_{t+1}^i)$ is continuous with respect to x_t^i, set U_t' is compact. Define set $U_t := \{u \in \mathbb{R}^I : u \leq u_t(\hat{y}_t), \text{ for some } \hat{y}_t \in \Gamma_t(\hat{\alpha}_{t+1})\}$, which by Assumption 3 is convex. Moreover, $U_t := U_t' - \mathbb{R}_+^I$. Compactness of U_t' implies that set U_t must be closed and bounded above.

Denote $u_t^* = u_t(\hat{x}_t)$. By definition, $\hat{x}_1$ is post-t efficient, so it must be $u_t^* \in \partial U_t$. By the Separating Hyperplane Theorem, there exists some non-zero vector $\alpha_t \in \mathbb{R}^I$ such that, for all $u \in U_t$, we have $\alpha_t \cdot u_t^* \geq \alpha_t \cdot u$. Since $U_t - \mathbb{R}_+^I \subset U_t$, it must be that $\alpha_t \in \mathbb{R}_+^I$. Moreover, by construction, we have $\hat{x}_t \in \Psi_t(\hat{\alpha}_t)$, where $\hat{\alpha}_t := (\alpha_t, \hat{\alpha}_{t+1})$.

[13] By $\geq$ we denote the coordinate-wise order on $\mathbb{R}^I$.

Finally, as for all $\hat{x}_t^i \in \hat{X}_t$, function $u_t^i(x_t^i, \hat{x}_{t+1}^i)$ is strictly concave in x_t^i and $\Gamma_t(\hat{\alpha}_{t+1})$ is convex, we have $\Psi_t(\hat{\alpha}_t) = \{\hat{x}_t\}$.

In order to prove the second part of the theorem, take any $t \in T$. Assume that for some $j \in I$, we have $\alpha_t^j = 0$. Let $y_t = (y_t^i)_{i \in I}$, where for all $i \neq j$, we have $y_t^i = x_t^i + 1/(I-1)x_t^j$ and $y_t^j = 0$. By assumption, x_t^j is non-zero, hence, for all $i \neq j$ we have $y_t^i > x_t^i$. Clearly, $(y_t, \hat{x}_{t+1}) \in \Gamma_t(\hat{\alpha}_{t+1})$. Since, for all $\hat{x}_{t+1}^i \in \hat{X}_{t+1}$, function $u_t^i(x_t^i, \hat{x}_{t+1}^i)$ is strictly increasing with respect to the first argument, we have $\sum_{i \in I} \alpha_t^i u_t^i(y_t^i, \hat{x}_{t+1}^i) > \sum_{i \in I} \alpha_t^i u_t^i(x_t^i, \hat{x}_{t+1}^i)$. This contradicts that $\hat{x}_t \in \Psi_t(\hat{\alpha}_t)$ as well as $\hat{x}_1 \in \Psi_1(\hat{\alpha}_1)$. □

Proposition 1.1 implies that it is possible to determine a wide class of time-consistently overall Pareto efficient allocations by solving the social welfare optimisation problem. On the other hand, Theorem 1.2 provides conditions under which every time-consistently overall Pareto efficient allocation can be represented by a solution to the same maximisation problem. Therefore, the two results show when the two notions are equivalent.

The proof of Theorem 1.2 relies strongly on the strict concavity assumption imposed on the utility functions. Once we weaken the condition to *weak* concavity, there might exist a time-consistently overall Pareto efficient allocation which cannot be represented via a solution to the recursive social welfare optimisation problem. For example, take $T = \{1, 2\}$ and assume that $\hat{x}_1 = (x_1, x_2)$ is a time-consistently overall Pareto efficient allocation path. Moreover, suppose that there exists a unique (up to a scalar) vector of positive weights α_2 such that $x_2 \in \Psi_2(\alpha_2)$. Once u_2^i is (weakly) concave, set $\Psi_2(\alpha_2)$ is convex and contains set $R_2(x_2) := \{y_2 \in E_2 : \text{for all } i \in I,\ u_2^i(y_2^i) = u_2^i(x_2^i)\}$, i.e., the set of period 2 allocations which are Pareto equivalent to x_2 with respect to period 2 selves. However, in general the two sets are not equal. Since

$$\Gamma_1(\alpha_2) := \{(y_1, y_2) \in E_1 : y_2 \in \Psi_2(\alpha_2)\} \supsetneq \{(y_1, y_2) \in E_1 : y_2 \in R_2(x_2)\},$$

there might exist some $\hat{y}_1 = (y_1, y_2)$ in $\Gamma_1(\alpha_2)$ such that y_2 is Pareto unordered relatively to x_2 with respect to period 2 selves, i.e., $y_2 \notin R_2(x_2)$, but $\hat{y}_1$ is Pareto preferred with respect to period 1 agents. In such cases, $\hat{x}_1$ would never be a solution to the recursive social welfare maximisation, just like in the following example.

Example 1.2. Consider a pure exchange economy with two consumers and two goods $j = 1, 2$. Hence, $I = \{1, 2\}$ and $T = \{1, 2\}$. Let $X_2 = \mathbb{R}^2_+$, with its elements denoted by $x^i_2 = (x^{i1}_2, x^{i2}_2)$. For all $i \in I$, let period 2 preferences $u^i_2 : X_2 \to \mathbb{R}$ be defined by

$$u^i_2(x^i_2) := x^{i1}_2 + x^{i2}_2.$$

On the other hand, let period 1 preferences $u^i_1 : X_2 \to \mathbb{R}$ be defined by

$$u^i_1(x^i_2) := \sqrt{x^{i1}_2} + \gamma^i \sqrt{x^{i2}_2},$$

where $\gamma^1 = 1$, $\gamma^2 = 3$. Hence, we assume that period 1 preferences are defined solely over period 2 consumption bundles. Eventually, let the total endowment in the economy be $\sum_{i\in I} e^i_2 = (1, 1)$.

Note that allocation $x_2 = (x^1_2, x^2_2) = ((x^{11}_2, x^{12}_2), (x^{21}_2, x^{22}_2)) = ((0.8, 0.2), (0.2, 0.8))$ is time-consistently overall Pareto efficient. Clearly, it is post-2 efficient. To see that it is also post-1 efficient, observe that the set of date 2 allocations that are Pareto equivalent to x_2 with respect to period 2 preferences is

$$R_2(x_2) = \{y_2 \in \mathbb{R}^4_+ : y^{i1}_2 + y^{i2}_2 = 1, \text{ for } i = 1, 2, \text{ and } y^{1j}_2 + y^{2j}_2 = 1, \text{ for } j = 1, 2\}.$$

Once we maximise the sum $u^1_1 + u^2_1$ over the above set, we obtain x_2. Clearly, there is no other allocation which improves the welfare of period 1 selves without making worst of at least one incarnation at time 2. Hence, x_2 is post-1 efficient, and so time-consistently overall Pareto efficient. However, there is no $\hat{\alpha}_1 := (\alpha_1, \alpha_2)$ such that the allocation is a solution to the corresponding recursive social welfare optimisation problem.

Observe that $x_2 \in \Gamma_1(\alpha_2)$ only for these weights $\alpha_2 = (\alpha^1_2, \alpha^2_2)$ for which $\alpha^1_2 = \alpha^2_2$. However, then $\Gamma_1(\alpha_2) = \{y_2 \in \mathbb{R}^4_+ : y^{1j}_2 + y^{2j}_2 = 1, \text{ for } j = 1, 2\}$, while the set of Pareto equivalent allocations to x_2 with respect to period 2 preferences is $R_2(x_2)$, defined as above. Therefore, $R_2(x_2) \subsetneq \Gamma_1(\alpha_2)$.

Assume that there exist some period 1 weights $\alpha_1 = (\alpha^1_1, \alpha^2_1)$ such that x_2 maximises the recursive social welfare. Given the weights, the allocation has to satisfy the following first order conditions:

$$\frac{\alpha^1_1}{\alpha^2_1} = \left(\frac{x^{11}_2}{x^{21}_2}\right)^{\frac{1}{2}}, \quad \text{and} \quad \frac{\alpha^1_1}{\alpha^2_1} = 3\left(\frac{x^{12}_2}{x^{22}_2}\right)^{\frac{1}{2}}.$$

However, since for the time-consistently overall Pareto efficient allocation x_2, we have $(x_2^{11}/x_2^{21})^{\frac{1}{2}} = 2$ and $3\,(x_2^{12}/x_2^{22})^{\frac{1}{2}} = 3/2$, there exists no such α_1 for which the above conditions are met.

Once we restrict our attention to strictly concave utility functions, set $\Psi_2(\alpha_2)$ is a singleton and the case discussed above does not occur.

1.5.3 Competitive equilibrium and social welfare

Combining Theorems 1.1 and 1.2 allows us to define a recursive social welfare optimisation problem with maximisers coinciding with any complete competitive equilibrium allocation. Consider the following proposition.

Proposition 1.2. *Let Assumptions 1, 2, and 3 be satisfied, and* $(\hat{x}_1^*, \hat{p}_1^*)$ *be a complete competitive equilibrium, where* $\hat{x}_1^* = (\hat{x}_1^{*i})_{i \in I}$ *and* $\hat{x}_1^{*i} = (x_t^{*i})_{t \in T}$. *There exists a non-zero path of weights* $\hat{\alpha}_1 := (\alpha_t)_{t \in T}$, *with* $\alpha_t \in \mathbb{R}_+^I$, *such that* $\Psi_1(\hat{\alpha}_1) = \{\hat{x}_1^*\}$. *In addition, if Assumption 4 is satisfied and* x_t^{*i} *is non-zero, for all* $i \in I$ *and* $t \in T$, *then* $\hat{\alpha}_1$ *is strictly positive, that is,* $\alpha_t \in \mathbb{R}_{++}^I$, *for all* $t \in T$.

Proof. Theorem 1.1 implies, that for any competitive equilibrium $(\hat{x}_1^*, \hat{p}_1^*)$, allocation path $\hat{x}_1^*$ is time-consistently overall Pareto efficient. By Theorem 1.2, any time-consistently overall Pareto efficient allocation maximises the recursive social welfare for some real, positive, non-zero weights $\hat{\alpha}_1$. In particular, this is true for $\hat{x}_1^*$. In addition, whenever Assumption 4 is satisfied and, for all $i \in I$ and $t \in T$, x_t^{*i} is non-zero, then $\hat{\alpha}_1$ is strictly positive. □

The above result states, that every allocation arising in a complete competitive equilibrium can be represented by a solution to the recursive social welfare optimisation problem, given some path of weights $\hat{\alpha}_1$. What is more, the proposition implies that there exists a method of aggregating preferences of agents with time-variant tastes and representing them by a single agent in the same class of preferences. Clearly, as it was mentioned before, the social planner in our problem faces a very similar time-inconsistency issue as every individual agent in the economy. Moreover, given the representation, we know that the resulting choice constitutes an allocation arising in some complete competitive equilibrium.

1.5.4 Efficiency, social welfare and hyperbolic discounting

In general, the notion of time-consistent overall Pareto efficiency does not coincide with date-1 Pareto efficiency, nor renegotiation-proofness. However, in some special cases the three concepts may be equivalent. We devote this section to one such prominent example.

Suppose that each period the corresponding commodity space is $X_t := \mathbb{R}^n_+$, for all $t \in T$. Moreover, the utility function of period t self of consumer $i \in I$ is defined as follows. For all $i \in I$, there exists some function $v^i : \mathbb{R}^n_+ \to \mathbb{R}$ and some numbers δ, $\gamma \in (0, 1)$ such that

$$u^i_t(\hat{x}^i_t) := v^i(x^i_t) + \gamma \sum_{s=t+1}^{T} \delta^{s-t} v^i(x^i_s). \tag{1.3}$$

In particular, note that the instantaneous utility functions v^i may differ across consumers, but the present-bias and long term discount factors γ and δ are common for all agents. We proceed with the following proposition.

Proposition 1.3. *Consider an economy with preferences defined as in* (1.3)*, where function v^i is strictly increasing, strictly concave and once continuously differentiable, for all $i \in I$. Then, a strictly positive allocation path $\hat{x}_1$ is date-1 Pareto efficient if and only if there exists a vector $\alpha^* \in \mathbb{R}^I_{++}$ such that $\hat{x}_1$ is a recursive social welfare allocation for weights $\hat{\alpha}_1 := (\alpha_t)_{t \in T}$, where $\alpha_t = \alpha^*$, for all $t \in T$. Moreover, $\hat{x}_1$ is time-consistently overall Pareto efficient and renegotiation-proof.*[14]

Proof. First, we prove ($\Rightarrow$). Let $\hat{x}_1$ be a strictly positive, date-1 Pareto efficient allocation path. By a well-known result (see Mas-Colell, Whinston, and Green, 1995, Proposition 16.E.2), there exists some $\alpha^* \in \mathbb{R}^I_+$ such that $\hat{x}_1 \in \operatorname{argmax}_{\hat{y}_1 \in E_1} \sum_{i \in I} \alpha^{*i} u^i_1(\hat{y}^i_1)$. Since v^i is strictly increasing, for all $i \in I$, we have $\alpha^{*i} > 0$. Moreover, $\hat{x}_1$ satisfies the following necessary and sufficient first order conditions, for all i, $j \in I$ and $t \in T$:

$$\begin{aligned} \alpha^{*i} \nabla v^i(x^i_t) &= \alpha^{*j} \nabla v^j(x^j_t), \\ \sum_{i \in I} x^i_t &= \sum_{i \in I} e^i_t. \end{aligned}$$

Define a path of weights $\hat{\alpha}_1 := (\alpha_t)_{t \in T}$ such that $\alpha_t = \alpha^*$, for all $t \in T$. Note, that the unique recursive social welfare allocation for the weights also has to satisfy the

[14] We say that $x \in \mathbb{R}^n$ is strictly positive, if $x \in \mathbb{R}^n_{++}$.

above first order conditions. Therefore, it must be that $\hat{x}_1 \in \Psi_1(\hat{\alpha}_1^i)$. Moreover, by Proposition 1.1, the allocation is time-consistently overall Pareto efficient.

Next, we show ($\Leftarrow$). Take any $\alpha^* \in \mathbb{R}^I_{++}$ and define $\hat{\alpha}_1 := (\alpha_t)_{t\in T}$ such that $\alpha_t = \alpha^*$, for all $t \in T$. Take some $\hat{x}_1 \in \Psi_1(\hat{\alpha}_1)$. Clearly, it satisfies the above first order conditions, which implies that $\hat{x}_1 \in \text{argmax}_{\hat{y}_1 \in E_1} \sum_{i\in I} \alpha^{*i} u_1^i(\hat{y}_1^i)$. Hence, by Mas-Colell, Whinston, and Green (1995, Proposition 16.E.2) is date-1 Pareto efficient. Again, by Proposition 1.1 the allocation must be time-consistently overall Pareto efficient.

Finally, we show that $\hat{x}_1$ is renegotiation-proof. Let R be the set of all time-consistently overall Pareto efficient allocations. Clearly, $\hat{x}_1 \in R \subseteq E_1$. Since $\hat{x}_1$ is date-1 Pareto efficient, there exists no other allocation $\hat{y}_1$ in E_1 (hence, in R) such that for all $i \in I$, $u_1^i(\hat{y}_1^i) \geq u_1^i(\hat{x}_1^i)$, and for some i, $u_1^i(\hat{y}_1^i) > u_1^i(\hat{x}_1^i)$. □

The above result is not entirely new. It has been already shown by Luttmer and Mariotti (2007, Proposition 1) that in the case of economies were agents are represented by time-separable preferences, the set of date-1 Pareto efficient allocations and the set of renegotiation-proof allocations coincide as long as the discount factors are identical for all consumers. We show, that additionally every element of the two sets is time-consistently overall Pareto efficient and maximises the recursive social welfare for a specific set of weights.

Proposition 1.3 crucially uses the assumption that discount factors are symmetric across consumers. Only then the first order conditions characterising the three notions of efficiency are equivalent. Moreover, the proposition above does not require the quasi-hyperbolic specification of discounting. In fact, as long as values of discount factors in each period are equal for all consumers, the claim of Proposition 1.3 remains true. In particular, this holds for the hyperbolic specification of discounting.

Finally, the above proposition does not imply that complete competitive equilibria in the discussed class of economies are efficient according to Definition 4(i). The result only states that allocations which are Pareto efficient with respect to the initial selves coincide with a class of time-consistently overall Pareto efficient ones, and maximise the recursive social welfare optimisation problem for some specific, time-invariant weights. In fact, Luttmer and Mariotti (2007, Proposition 3) show that in general such allocations do not arise in a competitive equilibrium.

1.6 Concluding remarks

In this chapter we characterized normative properties of competitive equilibria in economies with time-dependent preferences. We reintroduced the notion of time-consistent overall Pareto efficiency and showed that any equilibrium allocation path is efficient in this sense. This way, we extended the result by Herings and Rohde (2006, Theorem 30) to economies with a complete market structure. In addition, we complemented the analysis of Luttmer and Mariotti (2007) by establishing a general property of equilibrium outcomes in this class of models. Moreover, our result implies that the notion of renegotiation proofness, which might be considered as a benchmark for efficiency in an economy in which it is not possible to commit not to renegotiate, might be too strong to be achieved by a decentralised economy (see also Luttmer and Mariotti, 2007 for a discussion). Finally, due to Herings and Rohde (2008) our result can be extended to a wider class of economies with a more dynamic market structure.

Additionally, in this chapter we presented a way of representing both the time-consistently overall Pareto efficient allocations as well as equilibrium outcomes by a solution to a specific social welfare optimisation problem. In particular, this result allows us to aggregate time-dependent preferences of agents. Therefore, we established the existence of a sophisticated representative consumer in the discussed class of economies.

Chapter 2

Revealed time-preference

2.1 Introduction

Consider an experiment in which, in every trial, a consumer is presented with a finite set of pairs (m, t), consisting of a monetary payment $m \in \mathbb{R}_+$ and a time-delay $t \in \mathbb{N}$ at which the payment is delivered. The agent is allowed to choose exactly one option from the set. Suppose that we can observe both the set of feasible options, denoted by A, and the corresponding choice (m, t). Given a finite number of repetitions of the experiment, under what conditions can the choices of the consumer can be rationalised? In other words, when is it possible to determine a function $v : \mathbb{R}_+ \times \mathbb{N} \to \mathbb{R}$ such that, for any observable set of options A and the corresponding choice (m, t), we have

$$v(m, t) \geq v(n, s), \text{ for all } (n, s) \in A?$$

Clearly, without any additional conditions imposed on v, the above problem is trivial, as any constant function would rationalise an arbitrary set of observations. For this reason, given our setting, we focus on a class of functions which are strictly increasing with respect to monetary payments and strictly decreasing with time-delays. In particular, we are interested in preferences that are separable with respect to the two variables. That is, we discuss conditions under which the observable choices of agents can be supported by a utility function $v(m, t) := u(m)\gamma(t)$, where $u : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly increasing, while $\gamma : \mathbb{N} \to (0, 1]$ is strictly decreasing. For obvious reasons, we shall refer to γ as to a *discounting function*.

The separable specification of preferences seems to be especially important from the economic point of view. The discounted utility model plays a crucial role throughout the economic analysis and is widely accepted as a valid normative standard for public policies, as well as a descriptively accurate representation of the actual behaviour of economic agents. However, in the recent years an important question was raised concerning the form of the discounting function that reflects the actual time-preferences of consumers. In particular, alternative specifications of hyperbolic and quasi-hyperbolic discounting were proposed, which could explain various observations anomalous in the model of exponential discounted utility, formerly dominant in economics. See Frederick, Loewenstein, and O'Donoghue (2002) for a detailed discussion concerning this topic.

We propose an axiomatic characterisation of time-preference in a framework where the domain of choices is restricted to pairs of monetary payments and time-delays. Our prize-time set-up is similar to the one discussed in Fishburn and Rubinstein (1982), Ok and Masatlioglu (2007), or Noor (2011). However, unlike in those papers, we do not take the preference relation of an agent as a primitive, rather, we assume that the observer can monitor only a finite number of choices made by the consumer. This restriction significantly affects the conditions characterising time-preference. Since the observable choices induce only an incomplete preference ordering over the space of prize-time pairs (m, t), the question is how to extend the relation in a way that is consistent with a certain type of utility function. Whether this is possible or not determines if a given data set can be rationalised by a specific form of time-preference. The main motivation of this chapter is to establish the testable restrictions of various models of inter-temporal choice. In particular, we are interested in conditions that would allow us to distinguish between different specifications of the discounted utility model, including the hyperbolic, quasi-hyperbolic, and exponential.

We consider our framework to be particularly relevant from the perspective of empirical applications. There are numerous examples of experiments in which subjects are asked to choose between different monetary payments delivered with various time-delays. This includes an extensive list of studies presented by Frederick, Loewenstein, and O'Donoghue (2002, Table 1), as well as the works by Chabris, Laibson, Morris, Schuldt, and Taubinsky (2008, 2009), Andersen, Harrison, Lau, and Rutström (2008), Benhabib, Bisin, and Schotter (2010), or Dohmen, Falk, Huffman, and Sunde (2012). The design of the experiments allows us to apply our results directly to the sets of observations they generate.

We begin our discussion in Section 2.2, where we introduce the notation as well as some preliminary results. Throughout this chapter our axiomatic characterisation is imposed on the directly revealed preference relation induced by the set of observations. We say that a pair (m, t) is *directly revealed preferred* to (n, s), whenever there exists at least one observation of the experiment such that both options were available, i.e., they both belonged to the corresponding feasible set A, and (m, t) was chosen. We shall denote $(m, t)\,\mathcal{R}^*(n, s)$. The main difficulty in our problem is to determine a pre-order that extends the directly revealed relation to the whole domain of the prize-time pairs $\mathbb{R}_+ \times \mathbb{N}$. Moreover, we need to guarantee that the ordering can be represented by a utility function that possesses the desirable properties, in particular, monotonicity and separability.

First, we state the necessary and sufficient conditions under which the set of observations can be rationalised by a utility function $(m, t) \to v(m, t)$ that is strictly increasing with respect to monetary payments m and strictly decreasing in time-delays t. We define a partial order $\geq_X$ such that $(m, t) \geq_X (n, s)$ whenever $m \geq n$ and $t \leq s$, which is strict if at least one of the above inequalities is strict. We show that the set of observations can be rationalised in the above sense, whenever there is no sequence $\{(m^i, t^i)\}_{i=1}^n$ of pairs observed in the experiment such that every subsequent element dominates the preceding one with respect to $\mathcal{R}^*$ or $\geq_X$, and $(m^1, t^1) >_X (m^n, t^n)$. Therefore, we evoke a special case of *generalised cyclical consistency* discussed in Nishimura, Ok, and Quah (2013), as well as the Rationalisability Theorem II presented in the same paper.

The first main result of this chapter is presented in Section 2.3, where we concentrate on the axiomatic characterisation of preferences representable by a separable utility function $v(m, t) := u(m)\gamma(t)$. Clearly, as it is a special case of the previous representation, cyclical consistency is still a necessary condition, however, it is no longer sufficient. For this reason we introduce an alternative restriction called *dominance axiom*. Roughly speaking, the condition states that there exists no collection of directly revealed preference relations $(m, t)\,\mathcal{R}^*(n, s)$ in which the distribution of payments n, appearing in the inferior options, *first order stochastically dominates* the distribution of prizes m in the preferred pairs, while the distribution of time-delays t *first order stochastically dominates* the distribution of delays s.

Our approach to the characterisation of time-preference via the notion of stochastic dominance is novel. However, the tools we use to show the necessity and sufficiency of our axioms are similar to those applied in the classical literature on intuitive probability and additive plausibility (see, e.g., Kraft, Pratt, and Seidenberg, 1959 or Scott, 1964). In particular, the dominance axiom has a similar flavour to the *cancellation law* used extensively in this area of research. Moreover, our restriction describing the separable formulation of time-preference resembles the condition characterising the expected utility hypothesis, introduced by Border (1992). In his paper, Border concentrates on observable choices over sets of lotteries, and discusses conditions under which they can be rationalised by the expected utility model. The restriction of *ex-ante dominance*, that he proposes, hinges on a specific form of the first order stochastic dominance between the observable and an alternative, hypothetical choice function. Even though the question we consider, the framework we specify, as well as the tools we apply are substantially different from those used by Border, the main line of our argument is similar.

The second main theorem of this chapter, discussed in Section 2.4, concentrates on conditions which would allow us to characterise the discounting function γ more precisely. In particular, we provide the axiomatic characterisation of the *weakly present-biased* specification of γ, for which ratio $\gamma(t)/\gamma(t+1)$ is a decreasing function of t. Therefore, under our formulation, the relative discounting between any two dates diminishes as they become more distant in the future. Equivalently, this is to say that the function has a *log-convex* extension to the domain of real numbers. We consider this class to be especially important as it contains all the well-known specifications of discounting, including hyperbolic, quasi-hyperbolic, and exponential.

The condition characterising this class of time-preference is summarised by the *cumulative dominance axiom.* Our restriction requires that there exists no collection of directly revealed relations $(m,t)\,\mathcal{R}^*(n,s)$ such that the distribution of payments n in the inferior options *first order order stochastically dominates* the distribution of payments m appearing in the preferred pairs, while the distribution of time-delays t

second order stochastically dominates the distribution s. The condition is similar to the dominance axiom. However, as we require for the discounting function γ to be "log-convex", in order to make sure that the cumulative dominance axiom holds, we also need to consider samples in which the distributions of time-delays are ordered with respect to the second order stochastic dominance. Therefore, the cumulative dominance axiom is more restrictive, as we need to verify a larger class of collections of elements of the directly revealed preference relation while performing the test.

Finally, in the second part of Section 2.4, we draw our attention to an axiomatic characterisation of two specific examples of weakly present-biased discounting functions, namely, quasi-hyperbolic and exponential. The testable implications of the two specifications are similar, however, distinguishable. The essence of the two restrictions is summarised in the *strong cumulative dominance axiom.* Loosely speaking, the two specifications of time-preference require that there is no collection of directly revealed relations $(m, t)\mathcal{R}^*(n, s)$ such that the distribution of monetary payments n in the inferior options (n, s) first order stochastically dominates the analogous distribution of payments m, while the sum of time-delays t appearing in the superior prize-time pairs (m, t) is greater than the sum of delays s on the right hand side.

The results we present in this thesis are not the first attempt to axiomatise time-preference in a setting with a finite number of observations. Echenique, Imai, and Saito (2014) characterise various forms of the time-separable model of inter-temporal choice in a framework in which agents choose streams of a one-dimensional consumption good rather than prize-time pairs. In their setting, an observation consists of a consumption path selected by the subject and the corresponding prices of the commodity in the periods for which the choice is made. The authors specify both the necessary and sufficient conditions under which the set of observations can be rationalised by different forms of time-separable preference. What is crucial to their result, is the assumption concerning concavity of the instantaneous utility function. This allows the authors to constrain their attention to the implications of the first order conditions characterising the solution to the consumer optimisation problem. Therefore, the restrictions they discuss refer to the model of time-separable preferences *with* a concave instantaneous utility function. Our framework allows us to concentrate solely on the core implications of the discounted utility theory. We dispense the assumptions that are not crucial to the hypothesis and characterise these observable restrictions which are pivotal to this class of models. Nevertheless, as our set-up differs substantially from the one adopted by Echenique, Imai, and Saito, our results are not comparable.

2.2 Preliminaries

We begin the analysis with a formal specification of our framework. Let the *domain* over which the agents determine their choices be defined by $X := \mathbb{R}_+ \times \mathbb{N}$. Each element $x = (m, t)$ of the set consists of a monetary payoff $m \in \mathbb{R}_+$ and a time-delay $t \in \mathbb{N}$ at which the payment is delivered.

Let K be a finite set enumerating the subsequent trials (repetitions) of the experiment. In each trial $k \in K$, an agent is asked to choose one element from a finite *set of feasible options* $A_k \subset X$. An *experiment*, denoted by $\mathcal{E}$, is a collection of sets of feasible options,

$$\mathcal{E} := \{A_k\}_{k \in K}.$$

In every trial $k \in K$ of the experiment the subjects are obliged to choose exactly one element from the corresponding set of feasible options A_k. Therefore, an *observation* is an ordered pair (A_k, x_k), consisting of the set A_k and the option $x_k \in A_k$ chosen by the agent. Given this, the *set of observations* from the experiment is defined by a collection of the ordered pairs

$$\mathcal{O} := \{(A_k, x_k)\}_{k \in K},$$

where $x_k \in A_k$, for all $k \in K$. Note that our framework allows for the agents to make multiple choices from a single set A_k. However, each such choice has to be treated as a separate trial ("with replacement").[1] Finally, define the set of *observable options* by

$$\mathcal{A} := \bigcup_{k \in K} A_k.$$

[1]Note that our framework does not allow for consumers to choose several options simultaneously from a single set of feasible options, as the sequence in which the elements are chosen matters. Suppose that an agents chooses two elements x and y from some set A ("without replacement"). Then the corresponding observations are either (A, x) and $(A \backslash \{x\}, y)$, or (A, y) and $(A \backslash \{y\}, x)$, depending on which option was chosen first.

Hence, set $\mathcal{A}$ contains all the possible pairs of monetary payments and time-delays that the agent was offered at least once during the experiment. Clearly, we have $\mathcal{A} \subset X$. Moreover, since both K and A_k are finite, for every $k \in K$, so is $\mathcal{A}$.

Let the set of *observable payments* be given by

$$\mathcal{M} := \{m \in \mathbb{R}_+ : (m,t) \in \mathcal{A}\}.$$

Therefore, $\mathcal{M}$ is the set of all monetary prizes that appeared in at least one option during the experiment. Throughout this chapter we shall denote the cardinality of the set by $|\mathcal{M}|$, while $\underline{m} := \min \mathcal{M}$ and $\overline{m} := \max \mathcal{M}$. We define the set of *observable time-delays* by

$$\mathcal{T} := \{t \in \mathbb{N} : (m,t) \in \mathcal{A}\},$$

with its cardinality denoted by $|\mathcal{T}|$. Analogously, let the least and the greatest element of the set be denoted by $\underline{t} := \min \mathcal{T}$ and $\bar{t} := \max \mathcal{T}$. Finally, let $\bar{\mathcal{A}} := \mathcal{M} \times \mathcal{T}$.

2.2.1 Revealed preference relations and mixed-monotonicity

In the following section we discuss properties of preference relations induced by the set of observations $\mathcal{O}$. For any two elements x and y in $\mathcal{A}$, we say that x is *directly revealed preferred* to y, if in at least one trial of the experiment both options x and y were feasible but the agent decided to choose x rather than y. Formally, we will say that the pair (x,y) belongs to set $\mathcal{R}^*$ defined by

$$\mathcal{R}^* := \{(x,y) \in \mathcal{A} \times \mathcal{A} : \text{there exists } A \in \mathcal{E} \text{ such that } x,y \in A \text{ and } (A,x) \in \mathcal{O}\}.$$

For convenience, we will denote $x\,\mathcal{R}^* y$ instead of $(x,y) \in \mathcal{R}^*$. Whenever both (x,y) and (y,x) belong to $\mathcal{R}^*$ we will say that x and y are *directly revealed indifferent*. We denote the symmetric part of the relation by $\mathcal{I}^*$, i.e.,

$$\mathcal{I}^* := \{(x,y) \in \mathcal{A} \times \mathcal{A} : x\,\mathcal{R}^* y \text{ and } y\,\mathcal{R}^* x\}.$$

Similarly, we shall write $x\,\mathcal{I}^* y$ in place of $(x,y) \in \mathcal{I}^*$. Note that we do *not* define the strict counterpart of $\mathcal{R}^*$.

The main purpose of our analysis is to establish conditions under which the set of observations $\mathcal{O}$ can be rationalised by a specific form of utility function. Clearly, one of the necessary conditions for rationalisation is existence of a transitive closure of

$\mathcal{R}^*$ over $\mathcal{A}$. A complete, transitive, and reflexive pre-order $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ is *consistent* with the directly revealed preference relation $\mathcal{R}^*$ if for any two $x,\, y \in \mathcal{A}$, we have $x\,\mathcal{R}^* y \;\Rightarrow\; x\,\mathcal{R}\,y$, or equivalently $\mathcal{R}^* \subseteq \mathcal{R}$. We shall denote the strict component of $\mathcal{R}$ by $\mathcal{P}$, i.e.,

$$\mathcal{P} := \{(x, y) \in \mathcal{A} \times \mathcal{A} : x\,\mathcal{R}\,y \;\text{ and }\; \neg(y\,\mathcal{R}\,x)\}.$$

As previously, we shall write $x\,\mathcal{P}\,y$ instead of $(x, y) \in \mathcal{P}$. Finally, the symmetric part of $\mathcal{R}$ will be denoted by $\mathcal{I}$. Hence, $x\,\mathcal{I}\,y$ if and only if $x\,\mathcal{R}\,y$ and $y\,\mathcal{R}\,x$.

For the purposes of this thesis, we shall concentrate on a specific class of consistent pre-orders. Let $\geq_X$ denote a partial order on X such that, for any $x = (m, t)$ and $y = (n, s)$ in X, we have $x \geq_X y$ whenever $m \geq n$ and $t \leq s$. Moreover, the relation is strict, and denoted by $x >_X y$, if at least one of the above inequalities is strict. Loosely speaking, we will say that option x is greater than y with respect to $\geq_X$, if it offers a higher payment at a shorter delay. A pre-order $\mathcal{R}$ is *mixed-monotone*, whenever for any two elements x and y in $\mathcal{A}$, $x \geq_X y$ implies $x\,\mathcal{R}\,y$. In addition, if $x >_X y$ then $x\,\mathcal{P}\,y$. The definition suggests that whenever an agent is presented with two options such that one of them has a (weakly) higher payoff and a (weakly) shorter delay than the other one, then the former option should be preferred. Clearly, not every set of observations admits a consistent mixed-monotone pre-order. In the following section we discuss conditions under which there exists such an extension of the directly revealed preference relation.

2.2.2 Mixed-monotone rationalisation

Set $\mathcal{O}$ is *rationalisable* if there exists a function $v : X \to \mathbb{R}$, strictly increasing with respect to the partial order $\geq_X$,[2] such that for all $(A, x) \in \mathcal{O}$,

$$v(x) \geq v(y), \;\text{ for all } y \in A.$$

In the remainder of the chapter we focus on conditions under which the set of observations can be rationalised by a utility function v that strictly increases in the value of the monetary payment $m \in \mathbb{R}_+$ and strictly decreases with respect to the time-delay $t \in \mathbb{N}$. We begin by introducing the following axiom.

[2] That is, for any $x,\, y \in X$, whenever $x >_X y$ then $v(x) > v(y)$.

Axiom 2.1 (Cyclical consistency)**.** *Take any sequence* $\{x^i\}_{i=1}^n$ *in* $\mathcal{A}$ *such that* $x^{i+1}\mathcal{R}^*x^i$ *or* $x^{i+1} \geq_X x^i$*, for all* $i = 1, \ldots, n-1$*, and* $x^1 \geq_X x^n$*. Then it must be that* $x^1 = x^n$*.*

The above axiom is a special case of the *generalised cyclical consistency* condition formulated by Nishimura, Ok, and Quah (2013). It requires that whenever there exists a sequence of observable options such that every subsequent element is directly revealed preferred or greater (with respect to $\geq_X$) than the previous one, then it cannot be that the first element of the sequence is strictly greater than the ultimate one. Clearly, the violation of this condition excludes the existence of a consistent, mixed-monotone pre-order on $\mathcal{A}$. In fact, by Nishimura, Ok, and Quah (2013, Rationalisability Theorem I), cyclical consistency is also a sufficient condition for the existence of such a pre-order. In order to make our presentation more transparent, we discuss the following example.

Example 2.1. Consider the following directly revealed preference relation:

$$(5,3)\,\mathcal{R}^*(15,4),\;\; (15,2)\,\mathcal{R}^*(10,1),\;\; (15,1)\,\mathcal{R}^*(25,3),\;\; \text{and}\;\; (25,4)\,\mathcal{R}^*(20,2).$$

It is easy to check that the set of observations inducing the above relation is cyclically consistent. Given Nishimura, Ok, and Quah (2013, Rationalisability Theorem I), it is both necessary and sufficient to propose a consistent, mixed-monotone relation $\mathcal{R}$ defined over the observable options. For example

$$(15,1)\,\mathcal{I}\,(25,3)\,\mathcal{P}\,(25,4)\,\mathcal{I}\,(20,2)\,\mathcal{P}\,(15,2)\,\mathcal{I}\,(10,1)\,\mathcal{P}\,(5,3)\,\mathcal{P}\,(15,4).$$

Clearly, the relation is both consistent and mixed-monotone.

Proposition 2.1. *Set* $\mathcal{O}$ *is rationalisable if and only if it is cyclically consistent.*

Proposition 2.1 is a special case of the result by Nishimura, Ok, and Quah (2013, Rationalisability Theorem II), who establish the necessity and sufficiency of the generalised cyclical consistency condition for the existence of a utility function rationalising the choice data in a general class of partially ordered spaces. Unfortunately, their argument supporting the claim is not constructive, which implies that the only way of verifying whether $\mathcal{O}$ is rationalisable is by referring directly to the definition of cyclical consistency. Even though finite, this method may be highly inconvenient for applications, especially when the set of observations is large. For this reason, in Appendix

B, we present an alternative, constructive proof of Proposition 2.1, which introduces a more convenient method of verifying rationalisability of the set of observations $\mathcal{O}$ in our framework.

The necessity of cyclical consistency for rationalisation is straightforward. Clearly, for any function v rationalising $\mathcal{O}$, and any sequence $\{x^i\}_{i=1}^n$ specified as in the definition of the axiom, we have

$$v(x^n) \geq v(x^{n-1}) \geq \ldots \geq v(x^2) \geq v(x^1) \quad \text{and} \quad v(x^1) \geq v(x^n),$$

which can be satisfied only if $x_1 = x_n$. On the other hand, the "sufficiency" part of the proof is more demanding. We show the result in three steps. First, in Lemma B.1 we argue that cyclical consistency implies existence of a consistent, mixed-monotone pre-order $\mathcal{R}$ over $\mathcal{A}$. In Lemma B.2, we show that whenever such a pre-order exists, we can always find a sequence of real numbers $\{v_m^t\}_{(m,t)\in\bar{\mathcal{A}}}$ such that for any two options (m,t), $(n,s) \in \bar{\mathcal{A}}$, whenever $(m,t) >_X (n,s)$ or $(m,t)\,\mathcal{P}\,(n,s)$ then $v_m^t > v_n^s$, while $(m,t)\,\mathcal{I}\,(n,s)$ implies $v_m^t = v_n^s$. Finally, in Lemma B.3, we use any such sequence of numbers to construct a function rationalising the set of observations. Every step of our argument is constructive. Therefore, it presents a direct method of verifying whether an arbitrary set of observations is rationalisable.

2.3 Discounted utility rationalisation

In this section we present the first main theorem of this chapter. We say that set $\mathcal{O}$ is rationalisable by a *discounted utility* function whenever there is a strictly increasing instantaneous utility function $u : \mathbb{R}_+ \to \mathbb{R}_+$ and a strictly decreasing discounting function $\gamma : \mathbb{N} \to (0,1]$, with $\gamma(0) = 1$, such that $v(m,t) := u(m)\gamma(t)$ rationalises $\mathcal{O}$. Clearly, cyclical consistency is a necessary condition for this form of representation. However, it is no longer sufficient. The following section is devoted to determining both necessary and sufficient conditions which allow for such a representation.

2.3.1 Dominance axiom

A *sample* of the directly revealed preference relation $\mathcal{R}^*$ is a finite, indexed collection $\{(x^i, y^i)\}_{i\in I}$ of elements in $\mathcal{R}^*$, where we denote $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, for some m^i, $n^i \in \mathcal{M}$ and t^i, $s^i \in \mathcal{T}$. We allow for the samples to be generated "with replacement". That is, a single element of $\mathcal{R}^*$ may appear more than once in a sample.

Axiom 2.2 (Dominance axiom). *For any sample $\{(x^i, y^i)\}_{i\in I}$ of $\mathcal{R}^*$, where we denote $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, such that for any $m \in \mathcal{M}$ and $t \in \mathcal{T}$, we have*

$$|\{i \in I : m^i \leq m\}| \;\geq\; |\{i \in I : n^i \leq m\}| \;\; \text{and} \;\; |\{i \in I : t^i \leq t\}| \;\leq\; |\{i \in I : s^i \leq t\}|,$$

all of the above conditions hold with equality.

The above axiom requires that whenever there exists a sample such that the distribution of monetary payments n^i, in the inferior options y^i, first order stochastically dominates the corresponding distribution of payments m^i, in the preferred options x^i, while the distribution of time-delays t^i, appearing on the left hand side of $\mathcal{R}^*$, first order stochastically dominates the distribution of s^i, then both distributions of monetary payments and time-delays have to be equal. Therefore, the axiom is violated whenever there exists a sample $\{(x^i, y^i)\}_{i\in I}$ of the directly revealed preference relation, with $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, such that the distribution of n^i stochastically dominates the distribution of m^i, the distribution of t^i stochastically dominates the distribution of s^i, and at least one of the two relations is strict. In order to make our presentation more transparent, we discuss the following example.

Example 2.2. Consider the directly revealed preference relation analysed in Example 2.1. We claim that the set of observations inducing the relation fails to satisfy the dominance axiom. In order to show this, we need to find at least one sample which violates the condition specified in the definition of the axiom. Take $\mathcal{R}^*$. Clearly, the set is a sample of itself. Note that, given the support $\{5, 10, 15, 20, 25\}$, the distribution of payments in the preferred $x^i = (m^i, t^i)$ and the inferior options $y^i = (n^i, s^i)$ are respectively $(1, 0, 2, 0, 1)$ and $(0, 1, 1, 1, 1)$. Similarly, given the support $\{1, 2, 3, 4\}$, the distributions of the corresponding time-delays are both equal to $(1, 1, 1, 1)$. Therefore, there exists a sample of $\mathcal{R}^*$ for which the distribution of payments n^i *strictly* first order stochastically dominates the distribution of m^i, while the distributions of the time-delays are equal. Hence, we conclude that the set of observations $\mathcal{O}$ violates the dominance axiom.

The above example indicates that dominance is a stronger condition than cyclical consistency, as there exist sets of observations satisfying the latter but failing the former. On the other hand the dominance axiom implies cyclical consistency, as we will show in the remainder of this section.

Interestingly, the axiom is both a necessary and sufficient condition for the set of observations $\mathcal{O}$ to be rationalisable by a discounted utility function. We summarise this result in the following theorem. The proof is presented in Appendix B.

Theorem 2.1. *Set of observations $\mathcal{O}$ is rationalisable by a discounted utility function if and only if it obeys the dominance axiom.*

In order to understand why the dominance axiom is a necessary condition, suppose that $\mathcal{O}$ is rationalisable by a discounted utility function $v(m,t) := u(m)\gamma(t)$. This implies that the set is at the same time rationalisable by function $w(m,t) := \phi(m)+\varphi(t)$, where $\phi := \log(u)$ and $\varphi := \log(\gamma)$. Moreover, under this transformation functions ϕ and φ preserve the strict monotonicity of u and γ respectively.

Take any sample $\{(x^i, y^i)\}_{i\in I}$ of the directly revealed preference relation $\mathcal{R}^*$, where $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$. By the definition of rationalisation, for any element of the sample, we have $\phi(m^i) + \varphi(t^i) \geq \phi(n^i) + \varphi(s^i)$. In particular, once we sum up all the inequalities with respect to $i \in I$, we obtain

$$\sum_{i\in I} \phi(m^i) + \sum_{i\in I} \varphi(t^i) \geq \sum_{i\in I} \phi(n^i) + \sum_{i\in I} \varphi(s^i).$$

Suppose that the sample is specified as in the definition of the axiom. Then, the corresponding distribution of monetary payments n^i first order stochastically dominates the distribution of m^i. Since function ϕ is strictly increasing, it must be that $\sum_{i\in I} \phi(m^i) \leq \sum_{i\in I} \phi(n^i)$. On the other hand, we know that the distribution of time-delays t^i first order stochastically dominates the distribution of s^i. By monotonicity of φ, this implies that $\sum_{i\in I} \varphi(t^i) \leq \sum_{i\in I} \varphi(s^i)$. However, given the previous condition, the two inequalities may hold only if they are satisfied with equality. This requires that the distribution of monetary payments m^i is equal to the distribution of n^i, while the distribution of time-delays t^i coincides with the distribution of s^i.

The above argument highlights the form of consistency which is expected from a discounted utility maximiser. As it was shown above, whenever the set of observations is rationalisable by a discounted utility function, it can also be rationalised by an additive utility function $w(m,t) := \phi(m) + \varphi(t)$, where ϕ is strictly increasing and φ strictly decreasing. Take any sample $\{(x^i, y^i)\}_{i\in I}$ of $\mathcal{R}^*$, where $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, and construct two lotteries: one with the support corresponding to the preferred pairs (m^i, t^i) and probabilities equal to the frequencies with which they appear in the sample; the other one supported by options (n^i, s^i) and probabilities defined by the frequencies with which they appear in $\{(x^i, y^i)\}_{i\in I}$. Whenever the agent is an expected utility maximiser, his preference over the two lotteries should be consistent with his choices over individual pairs. Hence, the gamble supported by the superior options (m^i, t^i) should be preferred. On the other hand, due to separability and monotonicity of the utility function, any violation of the dominance axiom would imply that the consumer would rather choose the lottery supported by (n^i, s^i). However, such behaviour cannot be reconciled with the discounted utility maximisation.

Showing that the dominance axiom is a sufficient condition for this form of rationalisation is more demanding, hence, we place the proof in Appendix B. Nevertheless, we present the main observation used in our argument in the following proposition.

Proposition 2.2. *Set $\mathcal{O}$ obeys the dominance axiom if and only if there exists a strictly increasing sequence $\{\phi_m\}_{m\in\mathcal{M}}$ and a strictly decreasing sequence $\{\varphi_t\}_{t\in\mathcal{T}}$ of real numbers such that $(m,t)\,\mathcal{R}^*(n,s)$ implies $\phi_m + \varphi_t \geq \phi_n + \varphi_s$.*

This result is implied by Lemmas B.4 and B.5 in Appendix B, as well as the necessity of the dominance axiom for rationalisation by a discounted utility function. To support the above proposition we apply a variation of Farkas' Lemma, commonly known as *Motzkin's Rational Transposition*. Using the result, we show that the system of inequalities implied by the directly revealed preference relation fails to have a solution only if the set of observations violates the dominance axiom.

It is worth pointing out the importance of Proposition 2.2 for the applicability of Theorem 2.1. The result presents an alternative way of verifying whether the set of observations obeys the dominance axiom. In fact, the proposition states that the axiom is equivalent to the existence of a solution to a system of linear inequalities. Since such systems are in general solvable, i.e., there exist algorithms which allow to determine in a finite number of steps whether a given system has a solution or not, we find the alternative method of verifying the axiom to be much more convenient.

In Theorem 2.1 we establish the testable implications of the mixed-monotone discounted utility model. Given our framework, an important question is whether it is possible to test only for separability of the utility function. Unfortunately, the answer is "no". This follows from the fact that the set of observations does not induce strict directly revealed preference relations. Therefore, any separable function $v(m,t) := u(m)\gamma(t)$, where both u and γ are constant, trivially rationalises any set of observations. Hence, the monotonicity conditions we impose are required to evaluate any observable restrictions of the discounted utility model in our framework.

In the remaining sections of this chapter we determine conditions under which the class of preferences rationalising the data can be characterised more precisely. In particular, we focus on finer restrictions imposed on the form of the discounting function γ. Before we proceed with our analysis, we would like to discuss a family of experiments for which a narrower specification of time-preference is *never* possible.

2.3.2 Anchored experiments and identification

We say that an experiment $\mathcal{E}$ is *anchored*, whenever it consists solely of *binary* sets of feasible choices A, and there exists some $x^* \in X$ such that, for all $A \in \mathcal{E}$, we have $x^* \in A$. In other words, in each trial of an anchored experiment the subjects are asked to choose between one fixed option x^* and some other element in X. We shall denote $x^* = (m^*, t^*)$.

There are several notable examples of anchored experiments that were performed in the literature, including Kirby and Marakovic (1964), Coller and Williams (1999), or Kirby, Petry, and Bickel (1999), Harrison, Lau, and Williams (2002). Therefore, this class of experiments is relevant from the empirical point of view. An important advantage of these tests is that the choices the subjects face are relatively simple, which minimises the chance of errors made by the agents. Nevertheless, as we show in the next result, the simplicity of the experiment substantially reduces the informativeness of the observations it generates.

Proposition 2.3 (Indeterminacy). *For any anchored experiment the following statements are equivalent.*

(i) Set $\mathcal{O}$ is rationalisable.

(ii) Set $\mathcal{O}$ is rationalisable by a discounted utility function.

(iii) For any discounting function $\gamma : \mathbb{N} \to (0,1]$, with $\gamma(0) = 1$, there is a strictly increasing function $u : \mathbb{R}_+ \to \mathbb{R}_+$ such that $v(m,t) := u(m)\gamma(t)$ rationalises $\mathcal{O}$.

The above proposition states that, given observations from an anchored experiment, one can only determine if the choices of the subject are rationalisable according to the definition in Section 2.2.2. Therefore, the design of the experiment does not allow to verify whether the observable choices can be rationalised by a narrower class of preferences. In particular, Proposition 2.3(iii) implies that once the set of observations is rationalisable, it can also be rationalised by virtually *any* form of discounting. Hence, we consider this class of experiments to be rather weak, as the data they produce do not allow for a conclusive specification of time-preference explaining the observable choices.

The argument supporting the above claim is three-fold. In order to prove implication (i) $\Rightarrow$ (ii), we show that for any anchored experiment, whenever the set of observations is cyclically consistent, it may violate the dominance axiom only if it contains an infinite number of elements. Clearly, by definition this can never hold. Implication (ii) $\Rightarrow$ (iii) follows from the fact that any directly revealed preference relation can only be expressed with respect to the option (m^*, t^*). Therefore, given any discounting function γ, while constructing the utility function u we simply need to assign a value $u(m)$ to every element m of $\mathcal{M}$ that satisfies $u(m) \geq u(m^*)\gamma(t^*)/\gamma(t)$, whenever $(m,t)\,\mathcal{R}^*(m^*,t^*)$, and $u(m) \leq u(m^*)\gamma(t^*)/\gamma(t)$ otherwise, which is always possible. Implication (iii) $\Rightarrow$ (i) is obvious. The full proof of the result is presented in Appendix B.

2.4 Weakly present-biased rationalisation

In the previous section we have determined the necessary and sufficient conditions under which the set of observations can be rationalised by a discounted utility function. We devote the remainder of the chapter to determine restrictions which allow for a narrower characterisation of the discounting function γ. In particular, we focus on a class of preferences that exhibit some degree of present-bias. We will say that a strictly decreasing discounting function $\gamma : \mathbb{N} \to (0, 1]$, with $\gamma(0) = 1$, is *weakly present-biased*, whenever function $\vartheta : \mathbb{N} \to \mathbb{R}_{++}$,

$$\vartheta(t) := \frac{\gamma(t)}{\gamma(t+1)},$$

is decreasing. In other words, we require that the relative discounting between any two subsequent periods decreases as the two dates are further away in the future. Equivalently, this is to say that there exists a *log-convex* extension of function γ to the domain of the real numbers. Our interest in the above class of functions is primarily justified by the fact that it contains the most commonly used forms of discounting. In particular, the exponential, quasi-hyperbolic, and hyperbolic discounting models are included in this family.

2.4.1 Cumulative dominance axiom

In this section we present the second main result of this chapter. A set $\mathcal{O}$ is rationalisable by a *weakly present-biased discounted* utility function, whenever there exists a strictly increasing instantaneous utility function $u : \mathbb{R}_+ \to \mathbb{R}_+$ and a strictly decreasing, weakly present-biased discounting function $\gamma : \mathbb{N} \to (0, 1]$, with $\gamma(0) = 1$, such that $v(m, t) := u(m)\gamma(t)$ rationalises $\mathcal{O}$.

Axiom 2.3 (Cumulative dominance axiom)**.** *Take any sample* $\{(x^i, y^i)\}_{i\in I}$ *of* $\mathcal{R}^*$, *where* $x^i = (m^i, t^i)$ *and* $y^i = (n^i, s^i)$, *such that, for any* $m \in \mathcal{M}$ *and* $t \in \mathcal{T}$, *we have*

$$|\{i \in I : m^i \leq m\}| \geq |\{i \in I : n^i \leq m\}| \text{ and } \int_{\underline{t}}^{t} |\{i \in I : t^i \leq z\}| dz \leq \int_{\underline{t}}^{t} |\{i \in I : s^i \leq z\}| dz.$$

Then, for all $m \in \mathcal{M}$ *and* $t \in \mathcal{T}$, *we have* $|\{i \in I : m^i \leq m\}| = |\{i \in I : n^i \leq m\}|$ *and* $|\{i \in I : t^i \leq t\}| = |\{i \in I : s^i \leq t\}|$.

The cumulative dominance axiom requires that whenever there exists a sample such that the distribution of monetary payments n^i in the inferior options *first order stochastically dominates* the distribution of payments m^i in the preferred prize-time pairs, while the distribution of time-delays t^i appearing on the left hand side of $\mathcal{R}^*$

second order stochastically dominates the distribution of s^i, then the distributions of monetary payments and delays have to be equal. Roughly speaking, the axiom is violated whenever there exists at least one sample $\{(x^i, y^i)\}_{i \in I}$ of the directly revealed preference relation, with $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, such that the distribution of n^i first order stochastically dominates the distribution of m^i, and the distribution of t^i second order stochastically dominates the distribution of s^i, while at least one of the two relations is strict.

Note that the cumulative dominance axiom is a stronger requirement than the dominance axiom. Suppose that the set of observations satisfies the former condition. Then, there exists no sample of the directly revealed preference relation such that the monetary payments in the inferior options first order stochastically dominate the prizes in the preferred options, while the superior time-delays second order stochastically dominate the delays appearing on the right hand side of the relation. Since first order stochastic dominance implies the second, there exists no sample such the latter relation is preserved under the first order stochastic dominance. Hence, the set of observations must also satisfy the dominance axiom. However, the opposite implication does not hold, as there might exist a collection of elements in $\mathcal{R}^*$ such that the corresponding distributions of time-delays are not ordered with respect to the first order stochastic dominance, but are ordered with respect to the second. In other words, the cumulative dominance axiom requires verifying a larger set of samples than the dominance axiom.

Theorem 2.2. *Set $\mathcal{O}$ is rationalisable by a weakly present-biased discounted utility function if and only if it obeys the cumulative dominance axiom.*

In order to show the necessity of the cumulative dominance axiom for the narrower form of rationalisation, suppose that the choices of an agent can be explained by some function $v(m, t) := u(m)\gamma(t)$, where u is strictly increasing, while γ is strictly decreasing and weakly present-biased. Clearly, one can always rationalise the same set of observations by function $w(m, t) := \phi(m) + \varphi(t)$, with $\phi := \log(u)$ and $\varphi := \log(\gamma)$. The transformation preserves the monotonicity of the two functions, while φ has a convex extension to $\mathbb{R}_+$.

By definition, for any sample $\{(x^i, y^i)\}_{i\in I}$ of $\mathcal{R}^*$, where we denote $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, we have $\phi(m^i) + \varphi(t^i) \geq \phi(n^i) + \varphi(s^i)$, for all $i \in I$. In particular,

$$\sum_{i\in I} \phi(m^i) + \sum_{i\in I} \varphi(t^i) \geq \sum_{i\in I} \phi(n^i) + \sum_{i\in I} \varphi(s^i).$$

Suppose that the sample is specified as in the definition of the cumulative dominance axiom. As ϕ is strictly increasing, we have $\sum_{i\in I} \phi(m^i) \leq \sum_{i\in I} \phi(n^i)$. Moreover, since the distribution of time-delays t^i second order stochastically dominates the distribution of s^i, the existence of a convex extension of φ to the real space implies that $\sum_{i\in I} \varphi(t^i) \leq \sum_{i\in I} \varphi(s^i)$. However, the two inequalities can be consistent with the initial condition only if they are satisfied with equality which, by strict monotonicity of ϕ and φ, requires that the distributions of monetary payments and time-delays on the two sides of the directly revealed preference relation are equivalent.

The form of consistency that has to be satisfied by a weakly present-biased discounted utility maximiser is similar to the one discussed in Section 2.3. Take any sample $\{(x^i, y^i)\}_{i\in I}$ of $\mathcal{R}^*$, where $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, and construct two lotteries: one with the support corresponding to the preferred pairs (m^i, t^i) and probabilities equal to the frequencies with which they appear in the sample; the other one supported by options (n^i, s^i) and probabilities defined by the frequencies with which they show up in $\{(x^i, y^i)\}_{i\in I}$. Clearly, by construction, any expected utility maximiser prefers the former gamble to the latter. Given the Bernoulli utility function $w(m, t) := \phi(m)+\varphi(t)$, where ϕ is strictly increasing while φ is strictly decreasing and "convex", whenever the cumulative axiom is violated, it would be possible to construct a sample such that the lottery over the inferior options would be preferred to the gamble over the superior prize-time pairs. However, this violates the form of consistency that is required by the weakly present-biased discounted utility maximisation.

The "sufficiency" part of the proof of Theorem 2.2 is more demanding. Similarly as in the case of discounted utility, our argument consists of two steps. First, in Lemma B.7 (see Appendix B) we show that once the set of observations satisfies the cumulative dominance axiom, there always exists a solution to a specific system of linear inequalities. In the second step, see Lemma B.8, we use the solution to the system of inequalities in order to construct an instantaneous utility function u and a log-convex discounting function γ that rationalise the data. We present the key observation made in the proof of the theorem in the following proposition.

Proposition 2.4. *Set of observations $\mathcal{O}$ obeys the cumulative dominance axiom if and only if there is a strictly increasing sequence $\{\phi_m\}_{m\in\mathcal{M}}$, a strictly decreasing sequence $\{\varphi_t\}_{t\in\mathcal{T}}$, and a strictly negative sequence $\{v_t\}_{t\in\mathcal{T}}$ of real numbers such that $(m,t)\,\mathcal{R}^*(n,s)$ implies $\phi_m+\varphi_t \geq \phi_n+\varphi_s$, and for all $t \in \mathcal{T}$, we have $\varphi_t+v_t(s-t) \leq \varphi_s$, for all $s \in \mathcal{T}$.*

It is worth mentioning that the proposition is important for the applicability of our main result, as it presents an alternative method of verifying the cumulative dominance axiom. Moreover, we consider it to be much more convenient than applying the definition of the axiom directly.

2.4.2 Quasi-hyperbolic discounting

In the following section we concentrate on quasi-hyperbolic discounting functions, which constitute a narrower class of weakly present-biased preferences. We say that the set of observations is rationalisable by a *quasi-hyperbolic discounted utility* function, whenever there exist a strictly increasing function $u : \mathbb{R}_+ \to \mathbb{R}_+$, numbers β, $\delta \in (0,1)$, and some time-delay $t^\circ \in \mathbb{N}$ such that $v(m,t) := u(m)\gamma(t)$ rationalises $\mathcal{O}$, where

$$\gamma(t) := \begin{cases} \beta^t\delta^t & \text{for } t < t^\circ, \\ \beta^{t^\circ}\delta^t & \text{otherwise.} \end{cases}$$

Note that our definition of a quasi-hyperbolic discounting function generalises the standard notion for which $t^\circ = 1$. By allowing for the “threshold” time-delay t° to vary, we are able to analyse a wider class of preferences. In particular, as t° denotes a time-delay which separates the dates perceived by the agent as “present” from those regarded as “future”, we allow in our test for this parameter to be determined endogenously.

Since in the case of a quasi-hyperbolic discounting function $\vartheta(t) := \gamma(t)/\gamma(t+1)$ takes the value of $(\beta\delta)^{-1}$, for $t < t^\circ - 1$, and δ^{-1} otherwise, the quasi-hyperbolic specification is weakly present-biased. Hence, any set of observations rationalisable by this more specific form of time-preference obeys the cumulative dominance axiom. However, it is no longer sufficient. In this section we discuss the condition that fully characterises this form of the discounted utility model.

Axiom 2.4 (Strong cumulative dominance axiom)**.** *There exists some $t' \in \mathcal{T}$ such that, for any sample $\{(x^i, y^i)\}_{i\in I}$ of $\mathcal{R}^*$, where $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, satisfying*

(i) $|\{i \in I : m^i \leq m\}| \;\geq\; |\{i \in I : n^i \leq m\}|$, *for all* $m \in \mathcal{M}$*;*

(ii) $\sum_{i\in I} t^i \geq \sum_{i\in I} s^i$*;*

(iii) $\sum_{i\in I} \min\{t^i, t'\} \geq \sum_{i\in I} \min\{s^i, t'\}$,

all the above conditions hold with equality.

The above requirement is stronger than cumulative dominance. Clearly, in order to verify whether the above axiom is not violated, we need to consider a wider class of samples of the directly revealed preference relation. As previously, the samples of interest need to satisfy condition (i). However, additionally, they have to obey restrictions (ii) and (iii), which impose a weaker requirement on the relation between the distribution of time-delays appearing on both sides of the directly revealed preference relation in the sample. Therefore, a greater number of samples satisfies the two conditions then in the case of the cumulative dominance axiom.

Proposition 2.5. *Set $\mathcal{O}$ is rationalisable by a quasi-hyperbolic discounted utility function if and only if it obeys the strong cumulative dominance axiom.*

The necessity of the axiom can be proven similarly as in the previous sections. Suppose that the set of observations is rationalisable by a quasi-hyperbolic discounting function $v(m, t) := u(m)\gamma(t)$, where γ is specified as at the beginning of this section for some β, δ in $(0, 1)$, and a time-delay t°. This implies that the set of observations is also rationalisable by function $w(m, t) := \phi(m) + \min\{t, t^\circ\}\hat{\beta} + t\hat{\delta}$, where $\phi := \log(u)$, $\hat{\beta} := \log(\beta)$, and $\hat{\delta} := \log(\delta)$. By definition, for any sample $\{(x^i, y^i)\}_{i\in I}$ of the directly revealed preference relation, with $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, we have

$$\sum_{i\in I} \phi(m^i) + \hat{\beta} \sum_{i\in I} \min\{t^i, t^\circ\} + \hat{\delta} \sum_{i\in I} t^i \geq \sum_{i\in I} \phi(n^i) + \hat{\beta} \sum_{i\in I} \min\{s^i, t^\circ\} + \hat{\delta} \sum_{i\in I} s^i.$$

If the sample is specified as in the strong cumulative dominance axiom, it must be that $\sum_{i\in I} \phi(m^i) \leq \sum_{i\in I} \phi(n^i)$, as well as $\hat{\beta} \sum_{i\in I} \min\{t^i, t^\circ\} \leq \hat{\beta} \sum_{i\in I} \min\{s^i, t^\circ\}$ and $\hat{\delta} \sum_{i\in I} t^i \leq \hat{\delta} \sum_{i\in I} s^i$, since $\hat{\beta}$ and $\hat{\delta}$ are strictly negative. However, these conditions can be consistent with the initial inequality only if they all hold with equality. The argument remains unchanged once we substitute t° with $t' := \min\{t \in \mathcal{T} : t \geq t^\circ\}$. This requires for the strong cumulative dominance axiom to be satisfied.

The "sufficiency" part of the proof is presented in Appendix B. Our argument is constructed around one important observation, which we summarise below.

Lemma 2.1. *Set $\mathcal{O}$ obeys the strong cumulative dominance axiom for some $t' \in \mathcal{T}$ if and only if there exists a strictly increasing sequence $\{\phi_m\}_{m\in\mathcal{M}}$ and numbers $\hat{\beta}$, $\hat{\delta} < 0$ such that $(m,t)\,\mathcal{R}^*\,(n,s)$ implies $\phi_m + \min\{t,t'\}\hat{\beta} + t\hat{\delta} \geq \phi_n + \min\{s,t'\}\hat{\beta} + s\hat{\delta}$.*

Proposition 2.5 requires some comment. First of all, observe that the value of the time-delay t' for which set $\mathcal{O}$ obeys the axiom, determines the empirical "kink" of the quasi-hyperbolic discounting function. In particular, this means that we do not assume prior to the test the "threshold" date which separates the perceived "present" from the "future", but determine it endogenously. In fact, it is possible that one set of observations admits various forms of quasi-hyperbolic discounting, not only with respect to the values of the discount factors β and δ, but also with respect to the pivotal time-delay t'.

Second of all, Lemma 2.1 proposes an alternative method of verifying whether the set of observations obeys the axiom. As in Propositions 2.2 and 2.4, it hinges on the existence of a solution to a system of linear inequalities conditional on t'. Since the set of the observable time-delays is finite, the test can be performed in a finite number of steps.

2.4.3 Exponential discounting

Finally, we draw our attention to the exponential discounting models. We say that set $\mathcal{O}$ is rationalisable by an *exponential discounted utility* function whenever there is a strictly increasing instantaneous utility function $u : \mathbb{R}_+ \to \mathbb{R}_+$ and some $\delta \in (0,1)$ such that $v(m,t) := \delta^t u(m)$ rationalises the set of observations.

Proposition 2.6. *Set $\mathcal{O}$ is rationalisable by an exponential discounted utility function if and only if for any* subset $\{(x^i,y^i)\}_{i\in I}$ *of $\mathcal{R}^*$, where $x^i = (m^i,t^i)$ and $y^i = (n^i,s^i)$, such that $|\{i \in I : m^i \leq m\}| \geq |\{i \in I : n^i \leq m\}|$, for all $m \in \mathcal{M}$, and $\sum_{i\in I} t^i \geq \sum_{i\in I} s^i$, all the above conditions hold with equality.*

The above propositions states the necessary and sufficient condition under which a set of observations can be rationalised by an exponential discounted utility function. Note that the requirement significantly resembles the one stated in the definition of the strong cumulative dominance axiom. In fact, the only distinguishable implications of the quasi-hyperbolic and exponential model are implied by condition (iii) stated in the definition of the axiom. Clearly, Proposition 2.6 imposes a stronger condition on the set of observations, as it admits a larger class of samples that might violate it. However, whenever the strong cumulative dominance axiom is satisfied for t' equal to the least or the greatest observable time-delay, then the two requirements are equivalent. Observe that, given an arbitrary sample $\{(x^i, y^i)\}_{i \in I}$ of the directly revealed preference relation, condition (iii) stated in the axiom is trivially satisfied whenever $t' = \underline{t}$. On the other hand, if $t' = \bar{t}$, then conditions (ii) and (iii) coincide. This implies, that in these extreme cases, it is impossible to distinguish between the quasi-hyperbolic and exponential rationalisation. We summarise the result in the following corollary.

Corollary 2.1. *Set $\mathcal{O}$ is rationalisable by an expected utility function if and only if it obeys the strong cumulative dominance axiom for $t' = \underline{t}$ or $t' = \bar{t}$.*

The condition stated in Proposition 2.6 differs additionally in one substantial aspect from the strong cumulative dominance axiom. Observe that in order to rationalise the set of observations by an exponentially discounted utility function, we need to verify the requirement stated in the proposition only for *subsets* of the directly revealed preference relation, and not samples. Clearly, this substantially simplifies the test and reduces the number of steps required to test the condition.

2.4.4 Discount factors indeterminacy

In the following section we discuss some indeterminacy issues that arise while rationalising the set of observations by a quasi-hyperbolic or exponential discounted utility functions. We begin with the following proposition.

Proposition 2.7. *Set of observations $\mathcal{O}$ is rationalisable by a discounted utility function $v(m, t) := u(m)\gamma(t)$ if and only if, for any $a > 0$, it is rationalisable by a discounted utility function $\hat{v}(m, t) := \hat{u}(m)\hat{\gamma}(t)$, where $\hat{u} := u^a$ and $\hat{\gamma} := \gamma^a$.*

We omit the proof. The above result simply states that whenever a set of observations is rationalisable by some discounted utility function, then it is also rationalisable by its positive exponential transformation. This observation is not new, and was already noted by Fishburn and Rubinstein (1982, Theorem 2) in their representation theorem. However, the above proposition implies a much stronger conclusion. Namely, any property of function γ that is preserved under positive exponential transformations, is also satisfied by function $\hat{\gamma}$. In particular, this means that whenever γ is weakly present-biased (respectively quasi-hyperbolic or exponential) then so is $\hat{\gamma}$. This plays an important role in the next result.

Corollary 2.2. *Set $\mathcal{O}$ is rationalisable by a quasi-hyperbolic discounted utility function if and only if for any β (or δ) in $(0,1)$ there is some δ (respectively β) in $(0,1)$, a strictly increasing utility function $u:\mathbb{R}_+\to\mathbb{R}_+$, and a time-delay t° such that $v(m,t):=\gamma(t)u(m)$ rationalises $\mathcal{O}$, where $\gamma(t)=\beta^t\delta^t$, for $t<t^\circ$, and $\gamma(t)=\beta^{t^\circ}\delta^t$ otherwise.*

Proof. We show ($\Rightarrow$). There exists some strictly increasing function $u:\mathbb{R}_+\to\mathbb{R}_+$, discount factors β, δ in $(0,1)$, as well as time-delay t° such that $v(m,t):=u(m)\gamma(t)$ rationalises $\mathcal{O}$, where function γ is defined as at the beginning of this section. Take any $\beta'\in(0,1)$, and define $a:=\log(\beta')/\log(\beta)$. Clearly, function $\hat{v}(m,t):=\hat{u}(m)\hat{\gamma}(t)$, where $\hat{u}=u^a$ and $\hat{\gamma}=\gamma^a$, also rationalises $\mathcal{O}$. Moreover, $\hat{\gamma}(t)=(\beta')^t(\delta^a)^t$, for $t<t^\circ$, and $\hat{\gamma}(t)=(\beta')^{t^\circ}(\delta^a)^t$ otherwise. We present an analogous argument for the claim inside the brackets. Implication ($\Leftarrow$) is trivial. □

Given the nature of our framework and the above result, there is no testable restriction for the values of the discount factors β or δ, as long as we consider the two parameters separately. However, there exists a restriction for pairs (β,δ) of the two discount factors. In fact, a straightforward application of the above result allows to show that the restriction can be imposed on the ratio $\log(\delta)/\log(\beta)$.

On the other hand, there are no observable implications for the value of the discount factor δ rationalising the set of observations under exponential discounting. In fact, once the choice data can be rationalised for one value of the discount factor, it can be rationalised for virtually any other value $\delta\in(0,1)$. This observation is directly implied by Corollaries 2.1 and 2.2.

Corollary 2.3. *Set $\mathcal{O}$ is rationalisable by an exponential discounted utility function if and only if, for any $\delta\in(0,1)$, there exists a strictly increasing $u:\mathbb{R}_+\to\mathbb{R}_+$ such that $v(m,t):=\delta^t u(m)$ rationalises $\mathcal{O}$.*

Even though it is impossible to determine the value of the discount factor while rationalising the set of observations by an exponential discounted utility function, we are still able to impose the testable restriction on this class of models. Therefore, the conditions stated in Proposition 2.6 allow us to evaluate the shape of the discounting function, but not the parameter characterising it.

2.5 Concluding remarks

In this chapter we established the testable implications of the principal models of time-preference. Our results build a bridge between the decision-theoretical approach to consumer choice, that provides an axiomatic characterisation of different forms of behaviour, and the experimental work, that elicits the tastes of agents from a finite list of observable choices. In particular, our analysis determines the observable restrictions for the discounted utility model with different formulations of the discounting function. Moreover, we introduce an easy-to-apply test which allows to implement our results in an empirical analysis.

As we have shown in this chapter, in general, it is possible to distinguish between different models of time-preference in the prize-time framework that we consider. However, several indeterminacy issues may arise in this class of tests. First of all, separability alone is not a testable property of the utility function. Moreover, there are some substantial restrictions concerning identification of the discount factors in the quasi-hyperbolic and exponential discounted utility models. Finally, given observations generated in an anchored experiment, it is impossible to distinguish between any form of time-preference that we discussed in this chapter. Therefore, the design of the experiment strongly affects the informativeness of the observations it generates.

Chapter 3

Testing for production with complementarities*

3.1 Introduction

In this chapter, we discuss the testable restrictions for production technologies with complementarities. Suppose that we observe a list of choices of input factors x made by a single firm, given some commonly observed prices p. Under what conditions imposed on the set of observations can we rationalise the decisions of the firm by profit-maximising behaviour with production complementarities? In this chapter, we refer to the notion of complementarity introduced by Edgeworth, according to which two inputs are complements whenever an increase of one of the factors increases the marginal returns from the other one. This implies that whenever the price of one input falls, it is beneficial for the firm to increase the amount of both factors employed in production.

The importance of complementarities for "modern manufacturing" was highlighted by Milgrom and Roberts (1990, 1994, 1995), who observed the fundamental shift from mass production, which benefits from economies of scale, to a new pattern of manufacturing based on flexibility and economies of scope, that took place in the final decades of the twentieth century. This new paradigm relies on a system-wide and coordinated approach to production, where the "fit" of various attributes of technology plays a fundamental role. The mathematical representation of complementarity in terms of supermodular functions, and the one of lattice programming techniques generally, provides a way of formalising the intuitive ideas of synergy and system

*This chapter was written jointly with John K.-H. Quah.

effects. These tools allow for a rigorous analysis of the implications of technologies with complementarities without any additional, superfluous assumptions (see, e.g., Milgrom and Roberts, 1990, Milgrom and Shannon, 1994, or Topkis, 1995). Given the importance of the notion of complementarity and its mathematical formalisation, it is crucial to determine whether the hypothesis provides any observable restrictions that could be tested.

The main purpose of our analysis is to determine the necessary and sufficient conditions under which the set of observations, consisting of input-price pairs (x, p), can be rationalised by the profit-maximising behaviour with production complementarities. In other words, we characterise properties of the data set which guarantee that there exists a supermodular function f, defined over the set of all possible input vectors, such that for any observation (x, p), we have

$$f(x) - p \cdot x \geq f(y) - p \cdot y,$$

for any feasible input vector y. We interpret $f(x)$ as the revenue the firm receives when it uses factors x. Alternatively, whenever we consider a price-taking firm producing a single good, it is possible to normalise the price of the output to 1 and interpret f as a production function. In this chapter, we present an axiomatic characterisation of the profit-maximising behaviour with production complementarities. Additionally, we provide an easy-to-apply test which allows us either to confirm or refute the hypothesis on data sets.

Our revealed preference approach is motivated by the work of Afriat (1972), who studied the testable implications of the profit-maximisation hypothesis with an arbitrary production technology. Moreover, some of the issues that are closely related to our framework were also discussed by Rockafellar (1970), Rochet (1987), Brown and Calsamiglia (2007), and Sákovics (2013). Nevertheless, our analysis is substantially different from the one performed in the above papers. First of all, unlike Afriat, we assume that the output of the individual firms is not observable, and concentrate solely on their expenditure data. Second of all, we are not interested in rationalising the data set by an arbitrary production function. Rather, we determine the necessary and sufficient conditions that allow us to justify the observations with a production function exhibiting complementarities. In particular, we show that there are observable restrictions for this kind of technology. Recall that this is not true in the case of monotonicity or concavity of production (see, e.g., Sákovics, 2013). That is, unlike the above two characteristics, supermodularity of the production function is a testable property.

Ours is not the first attempt to characterise profit-maximisation with production complementarities. In fact, Chambers and Echenique (2009) formulate the so called *cyclical supermodularity* condition and show that it is both necessary and sufficient for the data set to be rationalisable by a supermodular production function. However, their results are formulated in the context of production functions defined on a *finite* sub-lattice of the Euclidean space. We extend their result by allowing the domain of the production function to be an arbitrary subset of $\mathbb{R}^\ell_+$. Moreover, we present a tighter condition which allows us to rationalise the set of observations by a supermodular *and* increasing production function. Hence, monotonicity of technologies with complementarities is a testable restriction in the framework we discuss.

Our analysis differs from the one performed by Chambers and Echenique in one additional aspect. Unlike these authors, who specify their axiom in terms of cycles of observable demands, we concentrate on their empirical distributions. By using the notion of *stochastic supermodular dominance*, introduced by Meyer and Strulovici (2013), we characterise data sets rationalisable by a supermodular production function through an intuitive condition on the distributions of observable demands generated by their finite collections drawn (possibly "with replacement") from the set of observations. Our approach is similar to the one applied in the classical literature on intuitive probability and additive plausibility (see, e.g., Kraft, Pratt, and Seidenberg, 1959 or Scott, 1964), as well as the one employed by Border (1992) to provide an axiomatic characterisation of the expected utility hypothesis. Moreover, it is closely related to the analysis presented in Chapter 2.

Finally, our results can be applied to rationalise consumer expenditure data by a quasilinear utility function, as in Brown and Calsamiglia (2007) or Sákovics (2013). We devote the final section of this chapter to discuss the similarities and differences of our approach to the one used in the related literature.

The remainder of the chapter is organised as follows. In Section 3.2 we present the key lattice-theoretical notions used in our analysis. In Section 3.3 we introduce our axioms as well as the main theorem of this chapter. Finally, Section 3.4 is devoted to the testable implications of the quasilinear utility-maximisation. The auxiliary results are presented in Appendix A which is placed at the end of this thesis.

3.2 Preliminaries

Let $\mathbb{R}^\ell$ be the real ℓ-dimensional space endowed with the standard, coordinate-wise partial order $\geq$. Hence, for any two real vectors $x = (x_i)_{i=1}^\ell$ and $y = (y_i)_{i=1}^\ell$ in $\mathbb{R}^\ell$, we have $x \geq y$ if and only if $x_i \geq y_i$, for all $i = 1, \ldots, \ell$. Moreover, for any x and y in $\mathbb{R}^\ell$, we denote their *join* and *meet* by $x \vee y$ and $x \wedge y$ respectively, where $(x \vee y)_i := \max\{x_i, y_i\}$ and $(x \wedge y)_i := \min\{x_i, y_i\}$, for $i = 1, \ldots, \ell$.

A pair $(X, \geq)$ is a sub-lattice of $(\mathbb{R}^\ell, \geq)$, whenever X is a subset of $\mathbb{R}^\ell_+$ and for any two elements x, $y \in X$, both $x \vee y$ and $x \wedge y$ belong to X.

Remark. With a slight abuse of the terminology, throughout this chapter we shall say that X is a *lattice* if and only if $(X, \geq)$ is a sub-lattice of $(\mathbb{R}^\ell, \geq)$.

For any lattices X and Y, we say that X dominates Y with respect to the *strong set order* if, for any $x \in X$ and $y \in Y$, we have $x \vee y \in X$ and $x \wedge y \in Y$. We will say that X is a *product lattice*, if it is defined by the Cartesian product of its projections on $\mathbb{R}$. Hence, $X := \times_{i=1}^\ell X_i$, where $X_i := \{y \in \mathbb{R} : y = x_i, \text{ while } (x_j)_{j=1}^\ell \in X\}$, for all $i = 1, \ldots, \ell$. Given any lattice X, we say that function $f : X \to \mathbb{R}$ is *supermodular*, whenever for any x and y in X, we have

$$f(x \vee y) + f(x \wedge y) \geq f(x) + f(y).$$

Suppose that lattice X is finite. For any vector $x \in X$, let $x + e_i$ denote the element y in X such that $y_j = x_j$, for all $j \neq i$, while $y_i := \min\{z \in X_i : z > x_i\}$, where X_i is the i'th projection of set X, $i = 1, \ldots, \ell$. In particular, note that e_i need not be a unit vector. For example, consider a finite lattice $X := \{(1, 2), (1, 4), (2, 2), (2, 4)\}$. Denote $x = (1, 2)$. Then, we have $x + e_1 + e_2 = (2, 2) + e_2 = (1, 4) + e_1 = (2, 4)$. Hence, even though e_1 is a unit vector, e_2 is not. Moreover, since X is a lattice, whenever $x + e_i$ and $x + e_j$ belong to X, for some $i \neq j$, then so does x and $x + e_i + e_j$. Finally, note that any vector e_i is defined conditionally on x.

Given any finite lattice X, by Topkis (1978, Theorems 3.1 and 3.2), function $f : X \to \mathbb{R}$ is supermodular if and only if, for any $x \in X$ such that $x + e_i$ and $x + e_j \in X$, $i \neq j$, we have

$$f(x + e_i + e_j) + f(x) \geq f(x + e_i) + f(x + e_j).$$

Moreover, function f is increasing whenever $f(x + e_i) \geq f(x)$, for any $x \in X$ and $i = 1, \ldots, \ell$ such that $x + e_i \in X$.

Throughout this chapter we find it convenient to refer to vectors $\varepsilon_x \in \{0, 1\}^{|X|}$, with the entry corresponding to element $x \in X$ equal to 1 and all the remaining coordinates equal to zero.

3.2.1 Supermodularity and stochastic dominance

In the following section we present the notion of *supermodular stochastic dominance* introduced in Meyer and Strulovici (2013). Throughout this section assume that X is a finite product lattice with cardinality $|X|$. We begin by defining a class of *elementary lattice transformations*. For any $x \in X$ such that $x + e_i + e_j$ belongs to X, $i \neq j$, let $t^x_{ij} : X \to \mathbb{R}^{|X|}$ denote a function defined by

$$t^x_{ij}(y) := \begin{cases} 1 & \text{if } \; y = x \;\text{ or }\; y = (x + e_i + e_j), \\ -1 & \text{if } \; y = (x + e_i) \;\text{ or }\; y = (x + e_j), \\ 0 & \text{otherwise.} \end{cases}$$

Equivalently, we may represent the elementary lattice transformation in terms of a vector $t^x_{ij} \in \{-1, 0, 1\}^{|X|}$ such that $t^x_{ij} := \varepsilon_{(x+e_i+e_j)} + \varepsilon_x - \varepsilon_{(x+e_i)} - \varepsilon_{(x+e_j)}$, where ε_x is defined as previously. We denote the set of all such functions by T.

Let Δ_X be the set of probability distributions over X. For any μ and $\nu \in \Delta_X$, we say that distribution μ dominates ν with respect to the *stochastic supermodular ordering*, and denote it by $\mu \succeq_S \nu$, if and only if for any supermodular function $f : X \to \mathbb{R}$, we have

$$\sum_{x \in X} f(x)\mu(x) \geq \sum_{x \in X} f(x)\nu(x). \tag{3.1}$$

Meyer and Strulovici (2013) present a dual characterisation of the above relation, which plays an important role in the remainder of this chapter.

Proposition 3.1. *For any two probability distributions μ, $\nu \in \Delta_X$, we have $\mu \succeq_S \nu$ if and only if there exist some non-negative coefficients $\{\lambda_t\}_{t\in T}$ such that*

$$\mu = \nu + \sum_{t\in T} \lambda_t t,$$

where T is the set of elementary lattice transformations on X.

See Meyer and Strulovici (2013) for the proof. Roughly speaking, we say that distribution μ dominates ν with respect to the stochastic supermodular dominance order, whenever the probability assigned by ν to any two unordered points x and y in its support is "shifted" in an equal value to $x \vee y$ and $x \wedge y$ in distribution μ. In other words, relatively to ν, distribution μ assigns an equally lower probability to any two unordered points x and y, and a respectively higher probability to their join $x \vee y$ and meet $x \wedge y$. We discuss the notion in the following example.

Example 3.1. Consider lattice $X := \{2, 4, 6\} \times \{1, 3, 5\}$ and two probability distributions μ and ν defined over X as follows. Distribution μ assigns probability $1/4$ to points $(2, 1)$, $(2, 3)$, $(4, 5)$, and $(6, 5)$, while ν assigns probability $1/4$ to points $(4, 1)$ and $(6, 3)$, and $1/2$ to $(2, 5)$. We depict the two distributions in Figure 3.1.

Note that μ dominates ν with respect to the stochastic supermodular dominance order. In fact, we can represent the former distribution by

$$\mu = \nu + \tfrac{1}{4}t_{12}^{(2,1)} + \tfrac{1}{2}t_{12}^{(2,3)} + \tfrac{1}{4}t_{12}^{(4,3)},$$

where the elementary lattice transformations t_{ij}^x are defined as previously. This is to say that we can construct μ by modifying distribution ν using only the elementary lattice transformations. In this case it suffices to shift the probability of $1/4$ from points $(2, 5)$ and $(6, 3)$ to their join and meet $(6, 5)$ and $(2, 3)$, and the probability of $1/4$ from points $(2, 5)$ and $(4, 1)$ to $(4, 5)$ and $(2, 1)$ respectively.

By definition of the elementary lattice transformations, for any two distributions μ and ν in Δ_X, if μ dominates ν with respect to the supermodular dominance order, then all marginal distributions of μ and ν are equal. In particular, whenever Y is the smallest product lattice such that $\text{supp}(\nu) \subseteq Y$, then for any probability distribution $\mu \in \Delta_X$, we have that $\mu \succeq_S \nu$ implies $\text{supp}(\mu) \subseteq Y$.

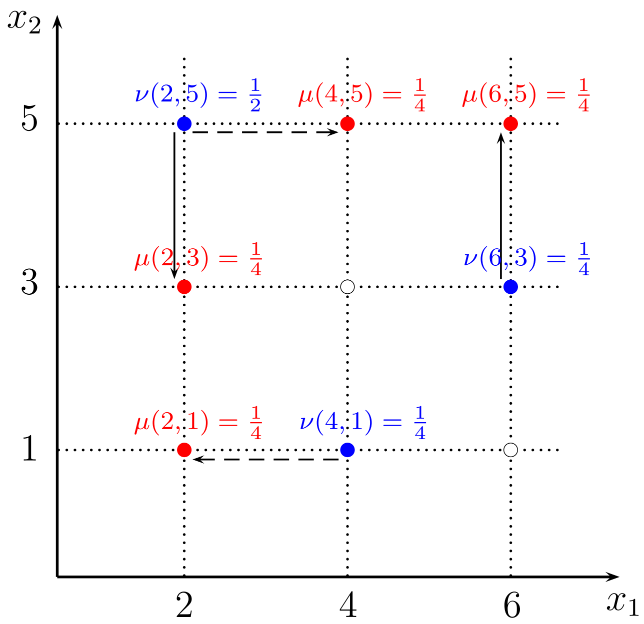

Figure 3.1: The figure depicts the distributions μ and ν discussed in Example 3.1. Observe that the probability of $1/4$ assigned to points $(2,5)$ and $(6,3)$ in ν is shifted equally to their join and meet, i.e., $(6,5)$ and $(2,3)$, in distribution μ (depicted by the solid arrows). An analogous shift takes place from points $(2,5)$ and $(4,1)$ to $(2,1)$ and $(4,5)$ (depicted by the dashed arrows). Hence $\mu \succeq_S \nu$.

3.2.2 Increasing supermodular dominance

The interest of this chapter also concerns increasing supermodular functions. First, we define the *elementary monotone transformations* on a finite product lattice X. For any $x \in X$ and $i = 1, \dots, \ell$ such that $x + e_i \in X$, define $\tau_i^x : X \to \mathbb{R}^{|X|}$ by

$$\tau_i^x(y) := \begin{cases} 1 & \text{if } \; y = x + e_i, \\ -1 & \text{if } \; y = x, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly to the elementary lattice transformations, we can represent function τ_i^x as a vector in $\{-1, 0, 1\}^{|X|}$ defined by $\tau_i^x := \varepsilon_{(x+e_i)} - \varepsilon_x$, where ε_x is specified as previously. Denote the set of all such transformations by $\mathcal{T}$.

For any two μ, $\nu \in \Delta_X$, we say that μ dominates ν with respect to the *increasing* stochastic supermodular ordering, and denote it by $\mu \succeq_{IS} \nu$, if and only if for any increasing and supermodular function $f : X \to \mathbb{R}$, condition (3.1) holds. As previously, Meyer and Strulovici present a dual characterisation of this notion.

Proposition 3.2. *For any two probability distributions μ, $\nu \in \Delta_X$, we have $\mu \succeq_{IS} \nu$ if and only if there exist some non-negative coefficients $\{\theta_\tau\}_{\tau\in\mathcal{T}}$ and $\{\lambda_t\}_{t\in T}$ such that*

$$\mu = \nu + \sum_{\tau\in\mathcal{T}} \theta_\tau \tau + \sum_{t\in T} \lambda_t t,$$

where $\mathcal{T}$ and T are the sets of the elementary monotone and lattice transformations on X respectively.

Roughly speaking, the increasing supermodular stochastic order is a combination of the supermodular and the first order stochastic dominance. Take any two probability distributions μ and ν such that $\mu \succeq_{IS} \nu$. Given Proposition 3.2, this implies that $\mu = \nu + \sum_{\tau\in\mathcal{T}} \theta_\tau \tau + \sum_{t\in T} \lambda_t t$ for some non-negative weights $\{\theta_\tau\}_{\tau\in\mathcal{T}}$ and $\{\lambda_t\}_{t\in T}$. Suppose that $\omega := \nu + \sum_{t\in T} \lambda_t t$ is a probability distribution. In particular, this implies that $\mu = \omega + \sum_{\tau\in\mathcal{T}} \theta_\tau \tau$. Hence, there exists a probability distribution ω such that $\omega \succeq_S \nu$, while μ first order stochastically dominates ω. Alternatively, whenever $\omega' := \nu + \sum_{\tau\in\mathcal{T}} \theta_\tau \tau$ is a probability distribution, then it first order stochastically dominates ν, while $\mu \succeq_S \omega'$ (see Meyer and Strulovici, 2013). The above description need not be precise, as in general ω and ω' may not be probability distributions.

Example 3.2. Consider the lattice discussed in Example 3.1. Suppose that μ assigns probability $1/4$ to points $(2,1)$, $(4,3)$, $(4,5)$, and $(6,5)$, while ν assigns probability $1/4$ to $(4,1)$ and $(6,3)$, and $1/2$ to $(2,5)$. Note that we can express μ by

$$\mu = \nu + \tfrac{1}{4}\tau_1^{(2,3)} + \tfrac{1}{4}t_{12}^{(2,1)} + \tfrac{1}{2}t_{12}^{(2,3)} + \tfrac{1}{4}t_{12}^{(4,3)},$$

where the monotone and lattice transformations τ_i^x and t_{ij}^x are defined as previously. Denote $\omega := \nu + \frac{1}{4}t_{12}^{(2,1)} + \frac{1}{2}t_{12}^{(2,3)} + \frac{1}{4}t_{12}^{(4,3)}$, which is a probability distribution that assigns probability $1/4$ to points $(2,1)$, $(2,3)$, $(4,5)$, and $(6,5)$. Hence, by Example 3.1, we have $\omega \succeq_S \nu$. Moreover, as $\mu = \omega + \frac{1}{4}\tau_1^{(2,3)}$, distribution μ first order stochastically dominates ω. Analogously, it is possible to construct a distribution $\omega' = \nu + \tau_1^{(2,3)}$ which first order stochastically dominates ν and $\mu \succeq_S \omega'$.

Similarly to the case of supermodular dominance, whenever distribution μ dominates ν with respect to the increasing supermodular dominance order, then any marginal distribution of μ first order stochastically dominates the corresponding marginal distribution of ν. In particular, there exist some product lattices Y_μ and Y_ν such that $\text{supp}(\mu) \subseteq Y_\mu$ and $\text{supp}(\nu) \subseteq Y_\nu$, while Y_μ dominates Y_ν with respect to the strong set order.

3.3 The main result

Suppose that we observe a finite set of price and factor input data which are elements of $\mathbb{R}^\ell_+$. Therefore, an observation is an ordered pair $(x, p) \in \mathbb{R}^\ell_+ \times \mathbb{R}^\ell_+$, consisting of an input vector x demanded at price p. Denote a *finite* set of observations by $\mathcal{O}$. The corresponding set of observable inputs is $\mathcal{X} := \{x \in \mathbb{R}^\ell_+ : (x, p) \in \mathcal{O} \text{ for some } p\}$. For any set $Y \subseteq \mathbb{R}^\ell_+$ such that $\mathcal{X} \subseteq Y$, we say that function $f : Y \to \mathbb{R}_+$ *rationalises* $\mathcal{O}$ *over* Y, if for any $(x, p) \in \mathcal{O}$, we have

$$f(x) - p \cdot x \geq f(y) - p \cdot y, \text{ for all } y \in Y.$$

We interpret $f(x)$ as the revenue the firm receives when is uses factors x. Alternatively, whenever we consider a price-taking firm producing a single good, it is possible to normalise the price of the output to 1 and interpret f as a production function.

In the remainder of this section we discuss the necessary and sufficient conditions on the set of observations which allow us to rationalise the data by profit-maximisation with production complementarities. In the following subsection we state two axioms characterising the set of observations, which will be central to our argument. Then, in Section 3.3.2, we present a simplified version of our analysis by constraining our attention to finite sub-lattices of $\mathbb{R}^\ell_+$. Finally, we state the main result of the chapter in Section 3.3.3, in which we extend the notion of rationalisation to an arbitrary sub-lattice of the real space.

3.3.1 Observable choices and supermodular dominance

Take any product lattice $Y \supseteq \mathcal{X}$ and define the following set of ordered triples:

$$\mathcal{R}_Y := \{(x, y, p) : (x, p) \in \mathcal{O} \text{ and } y \in Y\}.$$

Note that the above set captures the intuition of a directly revealed preference relation over elements in Y. Each triple, consists of two vectors of inputs x and y, and a vector of prices p, such that (x, p) constitutes a single observation, while y belongs to the product lattice Y. Clearly, under the assumption of profit maximisation, whenever we observe (x, p) it implies that, given prices p, acquiring inputs x yielded higher profit than y. Therefore, conditional on p, vector x was “preferred” by the firm to y. Moreover, we expect this condition to hold for any observation (x, p) in $\mathcal{O}$ and any vector $y \in Y$. In particular, whenever function $f : Y \to \mathbb{R}_+$ rationalises set $\mathcal{O}$ over Y, it must be that $(x, y, p) \in \mathcal{R}_Y$ implies $f(x) - p \cdot x \geq f(y) - p \cdot y$.

We define a *sample* of $\mathcal{R}_Y$ as a finite, indexed collection $\{(x^s, y^s, p^s)\}_{s\in S}$ of elements in $\mathcal{R}_Y$. In particular, we allow for a sample to be constructed "with replacement". That is, a single element of the set may appear more than once in the sample.

Axiom 3.1 (Supermodular dominance)**.** *For any sample $\{(x^s, y^s, p^s)\}_{s\in S}$ of $\mathcal{R}_Y$ define probability distributions μ and ν in Δ_Y by*

$$\nu(z) := \tfrac{1}{|S|}|\{s \in S : z = x^s\}| \quad \textit{and} \quad \mu(z) := \tfrac{1}{|S|}|\{s \in S : z = y^s\}|.$$

Set of observations $\mathcal{O}$ obeys the supermodular dominance axiom *over Y whenever for any sample of $\mathcal{R}_Y$, $\mu \succeq_S \nu$ implies $\sum_{s\in S} p^s \cdot (x^s - y^s) \leq 0$.*

The set of observations $\mathcal{O}$ obeys the above axiom over a product lattice $Y \supseteq \mathcal{X}$, whenever it is impossible to construct a sample $\{(x^s, y^s, p^s)\}_{s\in S}$ of $\mathcal{R}_Y$ such that the distribution of elements y^s dominates the distribution of x^s with respect to the supermodular stochastic ordering, while the total cost of inputs $\sum_{s\in S} p^s \cdot x^s$ is strictly greater than $\sum_{s\in S} p^s \cdot y^s$.

The above condition becomes more intuitive once we consider a set of observations consisting only of two elements (x^1, p^1) and (x^2, p^2). Consider a sample containing $(x^1, x^1 \vee x^2, p^1)$ and $(x^2, x^1 \wedge x^2, p^2)$. Whenever x^1 and x^2 are unordered, the distribution that assigns probability $1/2$ to $x^1 \vee x^2$ and $x^1 \wedge x^2$ dominates the distribution that assigns $1/2$ to x^1 and x^2 with respect to the stochastic supermodular ordering. Hence, for the axiom to hold, it must be that

$$p^1 \cdot x^1 + p^2 \cdot x^2 \leq p^1 \cdot (x^1 \vee x^2) + p^2 \cdot (x^1 \wedge x^2),$$

or equivalently, $p^2 \cdot (x^2 - x^1 \wedge x^2) \leq p^1 \cdot (x^1 \vee x^2 - x^1)$. That is, the difference between the cost of x^2 and $x^1 \wedge x^2$ under prices p^2 cannot be strictly greater than the difference in the cost of inputs $x^1 \vee x^2$ and x^1 under prices p^1. Once we consider sets with a greater number of observations, the above requirement becomes more sophisticated, as it has to be verified not only for individual pairs of input vectors, but also for their arbitrary sets. We discuss one such case in the following example.

Example 3.3. Consider lattice $Y = \{2,4,6\} \times \{1,3,5\}$. Suppose that we observe three input choices of a firm, equal to $(6,3)$, $(4,1)$, and $(2,5)$, at prices $(1,1)$, $(1.5,2.5)$, and $(2,2)$ respectively. The corresponding set $\mathcal{R}_Y$ consists of all triples (x,y,p), where x is equal to one of the observed choices, p is the price at which x was chosen, while y is any element in Y. We claim that the above observations fail to satisfy the supermodular dominance axiom. Consider the following sample of $\mathcal{R}_Y$:

$$\left\{\begin{array}{c}(x^1,y^1,p^1),\\(x^2,y^2,p^2),\\(x^3,y^3,p^3),\\(x^4,y^4,p^4)\end{array}\right\}=\left\{\begin{array}{c}((6,3),(4,5),(1,1)),\\((4,1),(2,3),(1.5,2.5)),\\((2,5),(2,1),(2,2)),\\((2,5),(6,5),(2,2))\end{array}\right\}.$$

Denote the distribution of x^s in the above collection by ν. Clearly, ν assigns probability 1/4 to points $(6,3)$ and $(4,1)$, and 1/2 to $(2,5)$. On the other hand, the distribution of y^s, which we denote by μ, assigns probability 1/4 to each element in its support, i.e., $(4,5)$, $(2,3)$, $(2,1)$, and $(6,5)$. By Example 3.1, we know that $\mu \succeq_S \nu$. Moreover, we have $\sum_{s=1}^4 p^s \cdot (x^s - y^s) = 1$, which is strictly greater than 0. Since there exists at least one sample violating the condition stated in the axiom, the set of observations fails to satisfy the supermodular dominance axiom.

Given the notion of increasing supermodular dominance order, we are able to specify a stronger requirement imposed on the set of observations.

Axiom 3.2 (Increasing supermodular dominance)**.** *For any sample $\{(x^s,y^s,p^s)\}_{s\in S}$ of $\mathcal{R}_Y$ define probability distributions μ and ν in Δ_Y by*

$$\nu(z) := \tfrac{1}{|S|}|\{s \in S : z = x^s\}| \quad \text{and} \quad \mu(z) := \tfrac{1}{|S|}|\{s \in S : z = y^s\}|.$$

Set of observations $\mathcal{O}$ obeys the increasing supermodular dominance axiom *over Y whenever for any sample of $\mathcal{R}_Y$, $\mu \succeq_{IS} \nu$ implies $\sum_{s\in S} p^s \cdot (x^s - y^s) \leq 0$.*

The above condition is similar to the supermodular dominance axiom, apart from the weaker notion of dominance used in the definition of the requirement. Clearly, the second axiom is stronger, as it requires to verify the condition for a larger class of samples of $\mathcal{R}_Y$. Namely, since $\mu \succeq_S \nu$ implies $\mu \succeq_{IS} \nu$, the set of samples that potentially violate the increasing supermodular dominance axiom is greater (by set inclusion) then the set of collections violating the supermodular dominance axiom. In the following example we present one set of observations that obeys the supermodular dominance axiom but fails to satisfy the increasing supermodular dominance axiom.

Example 3.4. Consider lattice Y discussed in Example 3.3. Suppose that the set of observations is $\mathcal{O} = \{(x^k, p^k)\}_{k=1}^4$, where $x^1 = x^2 = (2, 5)$, $x^3 = (4, 1)$, $x^4 = (6, 3)$, while $p^1 = (1, 2)$, $p^2 = (2, 2)$, $p^3 = (2, 1)$, and $p^4 = (3, 2)$. We do not verify it directly, but it is possible to show that the data set obeys the supermodular dominance axiom.[1] However, it does not satisfy the increasing supermodular dominance axiom. To show this, take the following sample of set $\mathcal{R}_Y$ induced by the set of observations:

$$\left\{\begin{array}{l}(x^1, y^1, p^1),\\(x^2, y^2, p^2),\\(x^3, y^3, p^3),\\(x^4, y^4, p^4)\end{array}\right\} = \left\{\begin{array}{l}((2,5),(2,1),(1,2)),\\((2,5),(4,5),(2,2)),\\((4,1),(6,5),(2,1)),\\((6,3),(4,3),(3,2))\end{array}\right\}.$$

By Example 3.2, we know that the distribution of y^s dominates the distribution of x^s with respect to the increasing supermodular dominance order. Moreover, we have $\sum_{s=1}^4 p^s \cdot (x^s - y^s) = 2$, which is strictly greater than 0. Hence, the increasing supermodular axiom is violated.

The axioms stated above have a similar flavour to the *cancellation law* used in the in the classical literature on intuitive probability and additive plausibility (see, e.g., Kraft, Pratt, and Seidenberg, 1959 or Scott, 1964). As in the above literature, our condition is imposed directly on the samples of elements drawn from the revealed preference relation. Clearly, the requirements are not equivalent, as the framework we discuss is different, however, the line of our argument highly resembles the one applied in those papers.

3.3.2 Supermodular rationalisation on a finite lattice

In the following section we show that the supermodular (increasing supermodular) dominance axiom is a necessary and sufficient condition for the observations to be rationalisable over a finite lattice by a supermodular (increasing and supermodular) technology. In particular, we revisit the result by Chambers and Echenique (2009) and show that the cyclical supermodularity requirement, specified in their paper, is equivalent to the supermodular dominance axiom. In the following proposition we

[1] Given Proposition 3.3, presented in the remainder of this chapter, we know that in order to verify the supermodular dominance axiom it is sufficient to determine a supermodular function $f : Y \to \mathbb{R}_+$ that rationalises the set of observations over Y. In fact, there exists at least one such function. For example, function $f(2,1) = 15$, $f(4,1) = 11$, $f(6,1) = f(6,5) = 1$, $f(2,3) = 13$, $f(4,3) = f(2,5) = 9$, and $f(6,3) = f(4,5) = 5$, is both supermodular and rationalises $\mathcal{O}$ over Y.

show the necessity of the supermodular (increasing supermodular) dominance axiom for rationalisation of the data.

Proposition 3.3. *Let $\mathcal{O}$ be rationalisable by a supermodular (increasing and supermodular) function $f : Y \to \mathbb{R}_+$ over Y, for some finite product lattice $Y \supseteq \mathcal{X}$. Then it obeys the supermodular (increasing supermodular) dominance axiom over Y.*

In order to show that the result in the brackets holds, take any increasing, supermodular function $f : Y \to \mathbb{R}_+$ and any sample $\{(x^s, y^s, p^s)\}_{s\in S}$ of $\mathcal{R}_Y$. By definition, for any $s \in S$, we have $f(x^s) - f(y^s) \geq p^s \cdot (x^s - y^s)$. In particular, once we sum up the inequalities with respect to $s \in S$, we obtain

$$\sum_{s\in S} f(x^s) - \sum_{s\in S} f(y^s) \geq \sum_{s\in S} p^s \cdot (x^s - y^s).$$

Given distributions ν and μ, defined as in the increasing supermodular dominance axiom, we have $\sum_{s\in S} f(x^s) = |S| \sum_{x\in Y} f(x)\nu(x)$ while $\sum_{s\in S} f(y^s) = |S| \sum_{x\in Y} f(x)\mu(x)$. Since $\mu \succeq_{IS} \nu$, by the definition of the increasing supermodular stochastic dominance,

$$0 \geq \sum_{s\in S} f(x^s) - \sum_{s\in S} f(y^s) \geq \sum_{s\in S} p^s \cdot (x^s - y^s).$$

Equivalently, we can show the necessity of the axiom using the dual representation of the increasing supermodular dominance order, specified in Proposition 3.2. Suppose that $\mu \succeq_{IS} \nu$. Therefore, there exist some positive weights $\{\theta_\tau\}_{\tau\in\mathcal{T}}$ and $\{\lambda_t\}_{t\in T}$ such that $\mu = \nu + \sum_{\tau\in\mathcal{T}} \theta_\tau \tau + \sum_{t\in T} \lambda_t t$, where $\mathcal{T}$ and T denote the sets of elementary monotone and lattice transformations respectively. In particular,

$$\begin{aligned}
\sum_{s\in S} f(x^s) - \sum_{s\in S} f(y^s) &= -|S| \sum_{x\in Y} f(x)(\mu(x) - \nu(x)) \\
&= -|S| \sum_{x\in Y} f(x) \left(\sum\nolimits_{\tau\in\mathcal{T}} \theta_\tau \tau(x) + \sum\nolimits_{t\in T} \lambda_t t(x)\right) \\
&= -|S| \sum_{\tau\in\mathcal{T}} \theta_\tau \sum_{x\in Y} f(x)\tau(x) - |S| \sum_{t\in T} \lambda_t \sum_{x\in Y} f(x)t(x).
\end{aligned}$$

Whenever function f is increasing, for any $\tau \in \mathcal{T}$ there is a $y \in Y$ with $y + e_i \in Y$ such that $\sum_{x\in Y} f(x)\tau(x) = f(y + e_i) - f(y) \geq 0$, for some $i = 1, \ldots, \ell$. Similarly, if the function is supermodular, then for any $t \in T$ there is a $y \in Y$ with $y + e_i + e_j \in Y$ such that $\sum_{x\in Y} f(x)t(x) = f(y + e_i + e_j) + f(y) - f(y + e_i) - f(y + e_j) \geq 0$, $i \neq j$.

Clearly, these two observations guarantee that $\sum_{s\in S} f(x^s) - \sum_{s\in S} f(y^s) \leq 0$, which by our previous argument implies $\sum_{s\in S} p^s \cdot (x^s - y^s) \leq 0$.

The result outside the brackets can be proven in an analogous way. Clearly, whenever $\mu \succeq_S \nu$, then for any supermodular function $f : Y \to \mathbb{R}_+$, we have $\sum_{s\in S} f(x^s) - \sum_{s\in S} f(y^s) \leq 0$. Hence, it must be that $\sum_{s\in S} p^s \cdot (x^s - y^s) \leq 0$. Similarly, it is possible to show the result using the dual representation of the stochastic supermodular dominance order. Recall that, whenever μ dominates ν in the above sense, we have $\mu = \nu + \sum_{\tau\in\mathcal{T}} \theta_\tau \tau + \sum_{t\in T} \lambda_t t$, where weights $\{\lambda_\tau\}_{\tau\in\mathcal{T}}$ are positive and $\theta_\tau = 0$, for all $\tau \in \mathcal{T}$. Hence, we can apply the same argument as previously.

Showing that the supermodular (increasing supermodular) dominance axiom is at the same time a sufficient condition for rationalisation is more demanding.

Proposition 3.4. *Suppose that $\mathcal{O}$ obeys the supermodular (increasing supermodular) dominance axiom over some finite product lattice $Y \supseteq \mathcal{X}$. There is a supermodular (increasing and supermodular) function $f : Y \to \mathbb{R}_+$ that rationalises $\mathcal{O}$ over Y.*

Proof. Enumerate the elements of $\mathcal{R}_Y$ such that the set is equal to $\{(x^r, y^r, p^r)\}_{r\in R}$. Let A be a $|\mathcal{T}| \times |Y|$ matrix, in which every row corresponds to a single elementary lattice transformation $\tau \in \mathcal{T}$. Therefore, for any $\tau_i^x \in \mathcal{T}$, the corresponding row is equal to $\varepsilon_{(x+e_i)} - \varepsilon_x$, where ε_x is defined as in Section 3.2. Similarly, let T be the set of the elementary lattice transformations defined over Y. By B we denote a $|T| \times |Y|$ matrix, in which every row corresponds to a single elementary lattice transformation $t \in T$. Hence, for any $t_{ij}^x \in T$, the corresponding row is equal to $\varepsilon_{(x+e_i+e_j)} + \varepsilon_x - \varepsilon_{(x+e_i)} - \varepsilon_{(x+e_j)}$. Let C be a $|R| \times |Y|$ matrix, where for each $r \in R$, the r'th row is equal to $\varepsilon_{x^r} - \varepsilon_{y^r}$. Finally, by d we denote a vertical vector of length $|R|$, where for each $r \in R$ the corresponding entry is equal to $p^r \cdot (x^r - y^r)$.

Observe that in order to prove the result outside the brackets, it suffices to show that, given the supermodular dominance axiom, there is a vector $f \in \mathbb{R}_+^{|Y|}$ such that $B \cdot f \geq 0$ and $C \cdot f \geq d$. On the other hand, proving the statement in the brackets is equivalent to showing that the increasing supermodular dominance axiom implies the existence of some $f \in \mathbb{R}_+^{|Y|}$ such that $A \cdot f \geq 0$, $B \cdot f \geq 0$, and $C \cdot f \geq d$. We prove only the second case. The result outside the brackets can be proven analogously.

First, we claim that whenever set $\mathcal{O}$ obeys the increasing supermodular dominance axiom, there is some $f \in \mathbb{R}^{|Y|}$ such that $A \cdot f \geq 0$, $B \cdot f \geq 0$, and $C \cdot f \geq d$. By

Theorem A.2 (see Appendix A), whenever the system of inequalities has no solution, there are vectors $\theta \in \mathbb{Z}_+^{|\mathcal{T}|}$, $\lambda \in \mathbb{Z}_+^{|T|}$ and $\gamma \in \mathbb{Z}_+^{|R|}$, where either $\theta > 0$, $\lambda > 0$, or $\gamma > 0$, such that $\theta \cdot A + \lambda \cdot B + \gamma \cdot C = 0$, while $\gamma \cdot d > 0$. Take any such vectors and denote $\theta = (\theta_\tau)_{\tau\in\mathcal{T}}$, $\lambda = (\lambda_t)_{t\in T}$, and $\gamma = (\gamma_r)_{r\in R}$. The above conditions can be rephrased as

$$\sum_{r\in R} \gamma_j \varepsilon_{x^r} + \sum_{\tau\in\mathcal{T}} \theta_\tau \tau + \sum_{t\in T} \lambda_t t = \sum_{r\in R} \gamma_r \varepsilon_{y^r} \quad \text{and} \quad \sum_{r\in R} \gamma_r p^r \cdot (x^r - y^r) > 0,$$

respectively. For any $x \in Y$, let $K(x) := \{r \in R : x = x^r\}$. That is, $K(x)$ is the set of all indices $r \in R$ enumerating the elements of $\mathcal{R}_Y$ for which the corresponding element x^r is equal to x. Analogously, let $L(y) := \{r \in R : y = y^r\}$. Define probability distributions ν and μ, by

$$\nu(x) := \frac{\sum_{k\in K(x)} \gamma_k}{\sum_{r\in R} \gamma_r} \quad \text{and} \quad \mu(y) := \frac{\sum_{l\in L(y)} \gamma_l}{\sum_{r\in R} \gamma_r}.$$ [2]

Therefore, probability $\nu(x)$ is equal to the sum of weights γ_r, $r \in R$, for which the corresponding element x^r is equal to x, divided by the sum of all γ_r. Analogously, probability $\mu(y)$ is equal to the normalised sum of weights γ_r, $r \in R$, for which the corresponding element y^r is equal to y. Moreover, denote $\hat{\theta}_\tau := \theta_\tau / \sum_{r\in R} \gamma_r$ and $\hat{\lambda}_t := \lambda_t / \sum_{r\in R} \gamma_r$. The above conditions imply that

$$\nu + \sum_{\tau\in\mathcal{T}} \hat{\theta}_\tau \tau + \sum_{t\in T} \hat{\lambda}_t t = \mu \quad \text{and} \quad \sum_{r\in R} \gamma_r p^r (x^r - y^r) > 0.$$

By Proposition 3.2, we have $\mu \succeq_{IS} \nu$, while $\sum_{r\in R} \gamma_r p^r (x^r - y^r) > 0$. Construct a sample $\{(x^s, y^s, p^s)\}_{s\in S}$ of $\mathcal{R}_Y$ such that each element (x^r, y^r, p^r) appears γ_r times, for all $r \in R$. Observe that the distributions of x^s and y^s are equal to ν and μ respectively. In addition, we have $\sum_{s\in S} p^s \cdot (x^s - y^s) > 0$. Since $\mu \succeq_{IS} \nu$, this violates the increasing supermodular dominance axiom. Therefore, we conclude that the axiom implies the existence of some $f \in \mathbb{R}^{|Y|}$ such that $A \cdot f \geq 0$, $B \cdot f \geq 0$, and $C \cdot f \geq d$.

In order to show that there is a solution in $\mathbb{R}_+^{|Y|}$, note that by definition of matrices A, B, and C, whenever vector f satisfies the above system of inequalities, then so does $f + a$, where a is a vector in $\mathbb{R}^{|Y|}$ with every entry equal to some constant. Clearly, we can always find an a such that $(f + a) \in \mathbb{R}_+^{|Y|}$. □

[2] In our notation we assume that $\sum_{i\in\emptyset} a_i = 0$.

Propositions 3.3 and 3.4 imply that the supermodular (increasing supermodular) dominance axiom is both a necessary and sufficient for the set of observations to be rationalisable by a supermodular (increasing and supermodular) function $f : Y \to \mathbb{R}_+$ over an arbitrary finite product lattice Y that contains $\mathcal{X}$. In particular, whenever we restrict our attention to product lattices, the result concerning the necessity and sufficiency of the supermodular dominance axiom for rationalisation of the data set is equivalent to Chambers and Echenique (2009, Proposition 2). This implies that the set of observations satisfies our axiom if and only if it is cyclically supermodular (see the above paper for a formal definition of this notion). In other words, we present an alternative characterisation of profit-maximisation with production complementarities in terms of distributions defined over the observable inputs, rather than their cycles. In addition, by introducing the increasing supermodular dominance axiom, we provide a stricter requirement under which the set of observations is rationalisable by an increasing and supermodular technology. Given Example 3.4, we know that monotonicity of production with complementarities is testable.

3.3.3 Supermodular rationalisation on an arbitrary lattice

In the following section we discuss conditions under which the set of observations can be rationalised over any sub-lattice of the Euclidean space. Before we state the main theorem, we need to introduce some additional notation. Let X^* denote the smallest (by set inclusion) product lattice that contains $\mathcal{X}$. In other words, we define $X^* := \times_{i=1}^{\ell} \mathcal{X}_i$, where $\mathcal{X}_i := \{y \in \mathbb{R} : y = x_i, \text{ where } (x_j)_{j=1}^{\ell} \in \mathcal{X}\}$ is the i'th projection of set $\mathcal{X}$. Consider the main theorem of this chapter.

Theorem 3.1. *Set $\mathcal{O}$ is rationalisable over $\mathbb{R}_+^{\ell}$ by a supermodular (increasing and supermodular) function $f : \mathbb{R}_+^{\ell} \to \mathbb{R}_+$ if and only if it obeys the supermodular (increasing supermodular) dominance axiom over X^*.*

The necessity part of the above theorem is directly implied by Proposition 3.3. Therefore, in the remainder of the section we focus on the sufficiency of the axioms. Let X_0^* be the smallest (by set inclusion) product lattice containing $\mathcal{X} \cup \{0\}$. In other words, we define $X_0^* := \times_{i=1}^{\ell} (\mathcal{X}_i \cup \{0\})$. Consider the following lemma.

Lemma 3.1. *Suppose that $\mathcal{O}$ obeys the supermodular (increasing supermodular) dominance axiom over X^*. Then it obeys the axiom over X_0^*.*

Proof. Take any sample $\{(x^s, y^s, p^s)\}_{s\in S}$ of $\mathcal{R}_{X_0^*}$ such that the distribution of y^s, denoted by μ, dominates the distribution of x^s, denoted by ν, with respect to the stochastic supermodular (increasing supermodular) dominance order. Denote the smallest (by set inclusion) product lattice containing supp(ν) by Y_ν. Similarly, let Y_μ denote the smallest product lattice containing supp(μ). Since $x^s \in \mathcal{X}$, for all $s \in S$, it must be that supp$(\nu) \subseteq X^*$. Moreover, if $\mu \succeq_S \nu$ then $Y_\mu = Y_\nu$, while $\mu \succeq_{IS} \nu$ implies that Y_μ dominates Y_ν with respect to the strong set order. In either case it must be that $Y_\mu \subseteq X^*$. Therefore, whenever $\mathcal{O}$ obeys the supermodular (increasing supermodular) dominance axiom over X^*, it satisfies the axiom over X_0^*. □

Given Proposition 3.4, the above lemma implies that whenever the set of observations obeys the supermodular (increasing supermodular) dominance axiom over X^*, there exists a supermodular (increasing and supermodular) function $f : X_0^* \to \mathbb{R}_+$, which rationalises the data over X_0^*. In the next lemma, we show that any such function has a supermodular (increasing and supermodular) extension to $\mathbb{R}_+^\ell$, which rationalises $\mathcal{O}$ over $\mathbb{R}_+^\ell$. Our argument is constructive. That is, we explicitly define the extension by applying a specific form of multi-linear interpolation between the points of $f(X_0^*)$. Before we prove the general result, we explain our method in following example.

Example 3.5. Consider two observations (x^1, p^1) and (x^2, p^2), where $x^1 = (3, 1)$, $p^1 = (2, 1)$, while $x^2 = (1, 2)$ and $p^2 = (1, 1)$. Hence, $\mathcal{X} = \{(3, 1), (1, 2)\}$, while the smallest product lattice containing $\mathcal{X}$ is $X^* = \{(1, 1), (3, 1), (1, 2), (3, 2)\}$. Finally, set X_0^* is equal to $X^* \cup \{(0, 0), (1, 0), (3, 0), (0, 1), (0, 2)\}$. In Figure 3.2 we plot the lattice as well as an increasing and supermodular function $f : X_0^* \to \mathbb{R}$ that rationalises the set of observations over X_0^*. We claim that there is an extension of f to $\mathbb{R}_+^2$.

Take any point in $\mathbb{R}_+^2$, e.g., $y = (y_1, y_2)$ depicted as in Figure 3.2. Note that the point belongs to the convex hull of elements $(1, 1)$, $(3, 1)$, $(1, 2)$, and $(3, 2)$. Clearly, it can be expressed as a linear combination of these points in the following way:

$$\begin{aligned} y &= (y_1, y_2) \\ &= \left(3 \cdot \tfrac{3-y_1}{3-1} + 1 \cdot \tfrac{y_1-1}{3-1},\ 2 \cdot \tfrac{2-y_2}{2-1} + 1 \cdot \tfrac{y_2-1}{2-1}\right) \\ &= \tfrac{2-y_2}{2-1} \cdot \left[\tfrac{3-y_1}{3-1} \cdot (3, 2) + \tfrac{y_1-1}{3-1} \cdot (1, 2)\right] + \tfrac{y_2-1}{2-1} \cdot \left[\tfrac{3-y_1}{3-1} \cdot (3, 1) + \tfrac{y_1-1}{3-1} \cdot (1, 1)\right] \\ &= \underbrace{\tfrac{2-y_2}{2-1} \cdot \tfrac{3-y_1}{3-1}}_{\alpha_1} \cdot (3, 2) + \underbrace{\tfrac{2-y_2}{2-1} \cdot \tfrac{y_1-1}{3-1}}_{\alpha_2} \cdot (1, 2) + \underbrace{\tfrac{y_2-1}{2-1} \cdot \tfrac{3-y_1}{3-1}}_{\alpha_3} \cdot (3, 1) + \underbrace{\tfrac{y_2-1}{2-1} \cdot \tfrac{y_1-1}{3-1}}_{\alpha_4} \cdot (1, 1). \end{aligned}$$

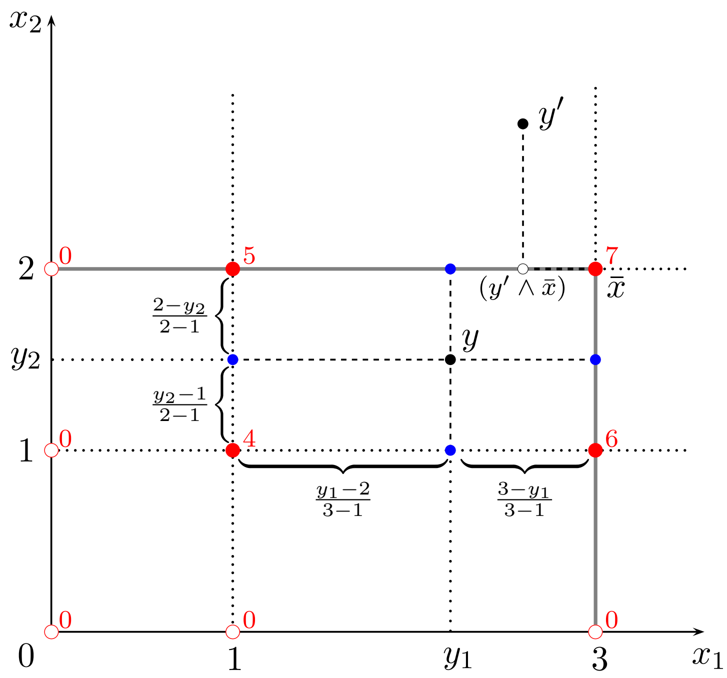

Figure 3.2: The figure depicts a production function rationalising the set of observations discussed in Example 3.5. The points in the corresponding lattice X^* are represented by solid red dots, while the void red dots represent the elements of $X_0^* \backslash X^*$. The values of function $f : X_0^* \to \mathbb{R}$ are denoted in red next to the corresponding points in X_0^*. The grey solid line denotes the upper bound of $\text{co}(X_0^*)$.

Given the above weights, we are able to define the value of function f at point y by taking a linear combination of its values at points in X^*. That is,

$$\begin{aligned} f(y) &= \alpha_1 f(3,2) + \alpha_2 f(1,2) + \alpha_3 f(3,1) + \alpha_4 f(1,1) \\ &= \alpha_1 \cdot 7 + \alpha_2 \cdot 5 + \alpha_3 \cdot 6 + \alpha_4 \cdot 4. \end{aligned}$$

Clearly, we can determine the value of the function in an analogous way for any point y that belongs to $\text{co}(X_0^*)$. Whenever a point is outside $\text{co}(X_0^*)$ (e.g., y' depicted in Figure 3.2) its value is equal to the value of the "closest" point that belongs to $\text{co}(X_0^*)$, i.e., $y' \wedge \bar{x}$, where $\bar{x} := \bigvee \text{co}(X_0^*) = (3,2)$.

In the following lemma we show that the extension of function $f : X_0^* \to \mathbb{R}_+$ to $\mathbb{R}_+^\ell$, constructed in the way presented above, preserves all the desired properties.

Lemma 3.2. *Let $f : X_0^* \to \mathbb{R}_+$ be a supermodular (increasing and supermodular) function that rationalises $\mathcal{O}$ over X_0^*. There exists a supermodular (increasing and supermodular) extension of f to $\mathbb{R}_+^\ell$ that rationalises set $\mathcal{O}$ over $\mathbb{R}_+^\ell$.*

Proof. Let $\text{co}(X_0^*)$ be the convex hull of X_0^*. Denote $I = \{1, \ldots, \ell\}$, and let $\mathcal{I}$ be the power set of I. Define $\mathring{X}_0^* := \{x \in X_0^* : x + \sum_{i\in I} e_i \in X_0^*\}$. In other words, $\mathring{X}_0^*$ consists of all the elements in X_0^* which are not contained in the upper bound of set. For every $x \in \mathring{X}_0^*$, define a hyperrectangle $H_x := \{y \in \mathbb{R}_+^\ell : x \leq y \leq x + \sum_{i\in I} e_i\}$, as well as the set of its vertices $V_x := \{y \in X_0^* : y = x + \sum_{i\in J} e_i,\text{ for some } J \in \mathcal{I}\}$. Observe that by definition, we have $H_x = \text{co}(V_x)$ and $X_0^* = \bigcup_{x\in\mathring{X}_0^*} V_x$. Clearly, this implies that

$$\text{co}(X_0^*) = \bigcup_{x\in\mathring{X}_0^*} H_x = \bigcup_{x\in\mathring{X}_0^*} \text{co}(V_x).$$

We denote the greatest element of $\text{co}(X_0^*)$ by $\bar{x}$.

Step 1: We propose a candidate function that extends f to $\mathbb{R}_+^\ell$. First, define correspondence $g : \text{co}(X_0^*) \rightrightarrows \mathbb{R}$ in the following way. For any element $x \in \mathring{X}_0^*$ and y in H_x, let $g(y) := \sum_{J\in\mathcal{I}} \alpha_J^x(y) f\left(x + \sum_{i\in J} e_i\right)$, where

$$\alpha_J^x(y) := \prod_{i\in I\setminus J} \frac{y_i - x_i}{\|e_i\|} \prod_{i\in J}\left(1 - \frac{y_i - x_i}{\|e_i\|}\right).[3]$$

and $\|\cdot\|$ is the Euclidean norm.[4] Therefore, for any $y \in H_x$, the value of $g(y)$ is a linear combination of the values of function f evaluated at the vertices of the hyperrectangle that contains y, i.e., points in $f(V_x)$. Moreover, the weights assigned to each element in $f(V_x)$ are defined such that $y = \sum_{J\in\mathcal{I}} \alpha_J^x(y)(x + \sum_{i\in J} e_i)$. In other words, we have

$$(y,\ f(y)) = \sum_{J\in\mathcal{I}} \alpha_J^x(y)\Big(x + \sum\nolimits_{i\in J} e_i,\ f\big(x + \sum\nolimits_{i\in J} e_i\big)\Big).$$

Finally, by definition of $\alpha_J^x(y)$, if y belongs to one of the sides (or edges, or is one of the vertices) of H_x, the strictly positive weights are assigned only to these vertices of the hyperrectangle that "span" that side (or that edge, or only that single vertex). In particular, this implies that g is single-valued at any non-empty intersection of hyperrectangles H_x. Since g is a linear function on the interior of each H_x, it is a continuous function over $\text{co}(X_0^*)$. Moreover, whenever f increases, so does g.

Define the extension of function f to $\mathbb{R}_+^\ell$, $\bar{f} : \mathbb{R}_+^\ell \to \mathbb{R}_+$, by $\bar{f}(y) := g(\bar{x} \wedge y)$. Clearly, for any $x \in X_0^*$, we have $\bar{f}(x) = g(x) = f(x)$. Hence, function $\bar{f}$ is a well

[3] In our notation, we assume that $\sum_{i\in\emptyset} a_i = 0$ and $\prod_{i\in\emptyset} a_i = 1$,

[4] Therefore, for any $i \in I$, $\|e_i\|$ is simply the distance between vectors x and $x + e_i$. More precisely, it is the difference between the values of the i'th coordinate of the two vectors.

defined extension. Moreover, as "$\wedge$" is a continuous operator and g is a continuous function, $\bar{f}$ is continuous. Finally, whenever f increases, so does $\bar{f}$.

Step 2: We show that function $\bar{f}$ is supermodular. We begin by showing that function g is supermodular over its domain. By Topkis (1978, Theorem 3.2), it suffices to show that for any $y_i' > y_i$, $y_j' > y_j$, $i \neq j$, we have

$$g(y_i', y_j', y_{-ij}) - g(y_i, y_j', y_{-ij}) \geq g(y_i', y_j, y_{-ij}) - g(y_i, y_j, y_{-ij}). \tag{3.2}$$

Since g is piece-wise linear, it is almost everywhere twice continuously differentiable. By the Fundamental Theorem of Calculus the above condition is equivalent to

$$\int_{y_j}^{y_j'} \int_{y_i}^{y_i'} \frac{\partial^2 g}{\partial z_i \partial z_j}(z_i, z_j, y_{-ij}) dz_i dz_j \geq 0.$$

For any $x \in \mathring{X}_0^*$, $i \in I$, and $y = (y_i, y_{-i})$ in the interior of H_x, we have

$$\frac{\partial g}{\partial y_i}(y_i, y_{-i}) = \sum_{J \in \mathcal{I}} \frac{\partial \alpha_J^x}{\partial (y_i)}(y) f\Big(x + \sum\nolimits_{k \in J} e_k\Big),$$

where for any $J \in \mathcal{I}$,

$$\frac{\partial \alpha_J^x}{\partial y_i}(y) := \begin{cases} \frac{1}{\|e_i\|} \prod\limits_{k \in J \setminus \{i\}} \frac{y_k - x_k}{\|e_k\|} \prod\limits_{k \in I \setminus J} \Big(1 - \frac{y_k - x_k}{\|e_k\|}\Big) & \text{if } i \in J, \\ -\frac{1}{\|e_i\|} \prod\limits_{k \in J} \frac{y_k - x_k}{\|e_k\|} \prod\limits_{k \in (I \setminus J) \setminus \{i\}} \Big(1 - \frac{y_k - x_k}{\|e_k\|}\Big) & \text{otherwise.} \end{cases}$$

Since $\mathcal{I}$ is the power set of I, for any set $J \in \mathcal{I}$ that contains i, there exists a corresponding set $J' = J \setminus \{i\}$. Equivalently, $(I \setminus J') \setminus \{i\} = I \setminus J$. Moreover, the above argument implies that $\partial \alpha_J^x / \partial y_i(y) = -\partial \alpha_{J'}^x / \partial y_i(y)$. Let $\mathcal{I}_i$ denote the collection of all subsets J of I such that $i \in J$. Then,

$$\frac{\partial g}{\partial y_i}(y) = \sum_{J \in \mathcal{I}_i} \frac{\partial \alpha_J^x}{\partial y_i}(y) \Big[f\Big(x + e_i + \sum\nolimits_{k \in J \setminus \{i\}} e_k\Big) - f\Big(x + \sum\nolimits_{k \in J \setminus \{i\}} e_k\Big)\Big].$$

For any $J \in \mathcal{I}_i$, denote

$$h(x, J) := f\Big(x + e_i + \sum\nolimits_{k \in J \setminus \{i\}} e_k\Big) - f\Big(x + \sum\nolimits_{k \in J \setminus \{i\}} e_k\Big).$$

Clearly, $\partial^2 g / \partial y_i \partial y_j(y) = \sum_{J \in \mathcal{I}_i} \partial^2 \alpha_J^x / \partial y_i \partial y_j(y) h(x, J)$, where, for any $J \in \mathcal{I}_i$,

$$\frac{\partial^2 \alpha_J^x}{\partial y_i \partial y_j}(y) := \begin{cases} \frac{1}{\|e_i\|\|e_j\|} \prod\limits_{k \in J \setminus \{j\}} \frac{y_k - x_k}{\|e_k\|} \prod\limits_{k \in I \setminus J} \Big(1 - \frac{y_k - x_k}{\|e_k\|}\Big) & \text{if } j \in J, \\ -\frac{1}{\|e_i\|\|e_j\|} \prod\limits_{k \in J} \frac{y_k - x_k}{\|e_k\|} \prod\limits_{k \in (I \setminus J) \setminus \{j\}} \Big(1 - \frac{y_k - x_k}{\|e_k\|}\Big) & \text{otherwise.} \end{cases}$$

Analogously, for any set $J \in \mathcal{I}_i$ that contains j, there exists some $J' = J\backslash\{j\}$, or equivalently $(I\backslash J')\backslash\{j\} = I\backslash J$. Moreover, $\partial^2\alpha^x_J/\partial y_i \partial y_j(y) = -\partial^2\alpha^x_{J'}/\partial y_i\partial y_j(y)$. Let $\mathcal{I}_{ij}$ denote the collection of all subsets J of I such that $\{i,j\} \subseteq J$. Therefore,

$$\frac{\partial^2 g}{\partial y_i \partial y_j}(y) = \sum_{J\in\mathcal{I}_{ij}} \frac{\partial^2 \alpha^x_J}{\partial y_i \partial y_j}(y)\big[h(x,J) - h(x,J\backslash\{j\})\big],$$

where $\partial^2\alpha^x_J/\partial y_i\partial y_j(y) \geq 0$. Moreover, for any $J \in \mathcal{I}_{ij}$, we have

$$h_i(x,J) - h_i(x,J\backslash\{j\}) = \big[f(z+e_i+e_j) - f(z+e_j)\big] - \big[f(z+e_i) - f(z)\big],$$

where $z = x + \sum_{k\in J\backslash\{j\}} e_k$. Since f is supermodular on X_0^*, the above expression is positive for all $z \in X_0^*$ such that $z + e_i + e_j \in X_0^*$. Hence, at any $y \in \text{co}(X_0^*)$ for which g is twice-continuously differentiable, we have $\partial^2 g/\partial y_i\partial y_j(y) \geq 0$. Therefore, condition (3.2) is always satisfied, which implies that function g is supermodular.

To show that $\bar{f}$ is also supermodular, take any y and z in $\mathbb{R}^\ell_+$. By definition of the function and supermodularity of g, we have

$$\begin{aligned}
\bar{f}(y) + \bar{f}(z) &= g(\bar{x}\wedge y) + g(\bar{x}\wedge z)\\
&\leq g\big((\bar{x}\wedge y)\vee(\bar{x}\wedge z)\big) + g\big((\bar{x}\wedge y)\wedge(\bar{x}\wedge z)\big)\\
&= g\big(\bar{x}\wedge(y\vee z)\big) + g\big(\bar{x}\wedge(y\wedge z)\big)\\
&= \bar{f}(y\vee z) + \bar{f}(y\wedge z).
\end{aligned}$$

Step 3: We show that function $\bar{f}$ rationalises $\mathcal{O}$ over $\mathbb{R}^\ell_+$. Take any $(x,p) \in \mathcal{O}$ and $y \in \mathbb{R}^\ell_+$. Since $\bar{x}\wedge y \in \text{co}(X_0)$, there exists some $x' \in \mathring{X}_0^*$ such that $\bar{x}\wedge y \in H_{x'}$. In particular, there exist weights $\{\alpha_k\}_{k=1}^{2^\ell}$, $\sum_{k=1}^{2^\ell}\alpha_k = 1$, such that $\bar{x}\wedge y = \sum_{k=1}^{2^\ell}\alpha_k z^k$, where $z^k \in V_{x'}$ is a vertex of hyperrectangle $H_{x'}$. By definition of function g, this implies that $g(\bar{x}\wedge y) = \sum_{k=1}^{2^\ell}\alpha_k f(z^k)$. Finally, f rationalises $\mathcal{O}$ over X_0^*. Hence, for any such x' and $z^k \in V_{x'}$, we have $f(x) - p\cdot x \geq f(z^k) - p\cdot z^k$. Therefore,

$$\begin{aligned}
\bar{f}(x) - p\cdot x &= f(x) - p\cdot x\\
&= \sum\nolimits_{k=1}^{2^\ell}\alpha_k f(x) - p\cdot\sum\nolimits_{k=1}^{2^\ell}\alpha_k x\\
&\geq \sum\nolimits_{k=1}^{2^\ell}\alpha_k f(z^k) - p\cdot\sum\nolimits_{k=1}^{2^\ell}\alpha_k z^k\\
&= g(\bar{x}\wedge y) - p\cdot(\bar{x}\wedge y)\\
&\geq g(\bar{x}\wedge y) - p\cdot y\\
&= \bar{f}(y) - p\cdot y,
\end{aligned}$$

where the first equality follows from the fact that $\bar{f}$ is an extension of f, and $x \in X^* \subseteq X_0^*$. The proof is complete. □

In some cases we might be interested in rationalising the set of observations by a production function which is defined over a subset of $\mathbb{R}_+^\ell$ (which might not necessarily be finite). In particular, this might take place when some input factors can only be acquired in integer amounts. The following corollary to Theorem 3.1 shows that the axioms stated in Section 3.3.1 remain both necessary and sufficient for this form of rationalisation.

Corollary 3.1. *Let Y be any lattice such that $X^* \subseteq Y$. Then, set $\mathcal{O}$ is rationalisable over Y by a supermodular (increasing and supermodular) function $f : Y \to \mathbb{R}$ if and only if it obeys the supermodular (increasing supermodular) dominance axiom over X^*. (Note that Y does not have to be a product lattice.)*

The above result follows from the fact that, whenever there exists a function rationalising $\mathcal{O}$ over $\mathbb{R}_+^\ell$, the same function rationalises the set over any restricted domain. Moreover, both supermodularity and monotonicity are preserved over the new domain. This implies that the supermodular (increasing supermodular) dominance axiom is a necessary and sufficient condition for rationalising the data set over an arbitrary sub-lattice of the Euclidean space.

Suppose that lattice Y is finite. The above observation implies that, whenever Y contains X^*, in order to rationalise $\mathcal{O}$ over Y, it is sufficient for the set of observations to obey the supermodular (increasing supermodular) dominance axiom over X^*, rather than Y. In other words, Corollary 3.1 provides a tighter condition than the result discussed in Propositions 3.3 and 3.4, as it suffices for our axioms to be verified over a smaller set. In particular, this also refers to Chambers and Echenique (2009, Proposition 2), who require for the cyclical supermodularity condition to be satisfied for any cycle in Y. Since their axiom is equivalent to supermodular dominance over Y, it is enough to check either the cyclical supermodularity condition or the supermodular dominance axiom for the elements of X^* *only*. This plays an important role for the applicability of our results to an empirical analysis. Clearly, whenever set Y is much larger than X^*, the above observation allows us to substantially reduce the complexity of the test.

Even though our axioms provide an intuitive condition for rationalisability of the set of observations, checking whether the requirement is satisfied by the data might be cumbersome and hard to apply in an empirical analysis. Fortunately, there exists a more convenient way of implementing the test. Consider the following proposition.

Proposition 3.5. *Let $Y \supseteq \mathcal{X}$ be a finite product lattice. Set $\mathcal{O}$ obeys the supermodular (increasing supermodular) dominance axiom over Y if and only if there is a sequence (increasing sequence) of numbers $\{f_x\}_{x\in Y}$ such that, for any $x \in Y$ with $x + e_i + e_j$ in Y, $i \neq j$, we have $f_{(x+e_i+e_j)} + f_x \geq f_{(x+e_i)} + f_{(x+e_j)}$, and $(x, y, p) \in \mathcal{R}_Y$ implies $f_x - p \cdot x \geq f_y - p \cdot y$.*

The result states that the supermodular (increasing supermodular) dominance axiom is equivalent to the existence of a solution to a certain system of linear inequalities. It is clear that any such solution constitutes a supermodular production function that rationalises set $\mathcal{O}$ over Y.[5] Therefore, Proposition 3.5 is simply a reformulation of Propositions 3.3 and 3.4. Note that the above result provides us with a simple method of verifying if $\mathcal{O}$ is rationalisable over $\mathbb{R}^\ell_+$. By Theorem 3.1, the set of observations is rationalisable over the Euclidean space if and only if it obeys the supermodular (increasing supermodular) dominance axiom over X^*. By Proposition 3.5, we can verify the condition by finding a solution to the corresponding system of linear inequalities. As it is well known, such systems can be efficiently solved in a finite number of steps, which we find crucial for the applicability of our result.

3.4 Related results and alternative applications

In the previous sections we were concentrating on the necessary and sufficient conditions under which the set of observations can be rationalised by either a supermodular or an increasing and supermodular production function. Clearly, the axioms we discussed are at the same time sufficient for rationalising $\mathcal{O}$ by *any* production function. However, as we were focusing on a specific type of technology, our conditions are not

[5] Obviously, we can define $f : Y \to \mathbb{R}$, by $f(x) := f_x$. Given this, for any $x \in Y$ with $x + e_i$ and $x+e_j$ in Y, $i \neq j$, we have $f(x+e_i+e_j)+f(x) \geq f(x+e_i)+f(x+e_j)$, which implies supermodularity of the function. Finally, as $(x, y, p) \in \mathcal{R}_Y$ implies $f(x) - p \cdot x \geq f(y) - p \cdot y$, function f rationalises $\mathcal{O}$ over Y.

necessary in this more general case. In the following section we show how our results relate to those in the existing literature.

In the second part of the section we discuss the way in which Theorem 3.1 corresponds to the testable restrictions for the quasilinear utility-maximisation hypothesis. In particular, we focus on how our results can be applied to the analysis performed by Brown and Calsamiglia (2007) or Sákovics (2013).

3.4.1 Cyclical monotonicity and supermodular dominance

The necessary and sufficient conditions for rationalisation of input-price data by an arbitrary technology can be obtained directly by the application of Rockafellar (1970, Theorem 24.8). Various versions of the argument can be found, e.g., in Brown and Calsamiglia (2007), Chambers and Echenique (2009), or Sákovics (2013). In order to make our thesis self-contained, we present the result below.

Proposition 3.6. *The following statements are equivalent.*

(i) Set $\mathcal{O}$ is rationalisable by a function $f : \mathbb{R}^\ell_+ \to \mathbb{R}_+$ over $\mathbb{R}^\ell_+$.

(ii) The set of observations is cyclically monotone*. That is, for any finite sequence $\{(x^n, p^n)\}_{n=1}^m$ in $\mathcal{O}$, we have*

$$p^1 \cdot (x^1 - x^2) + p^2 \cdot (x^2 - x^3) + \ldots + p^{m-1} \cdot (x^{m-1} - x^m) + p^m \cdot (x^m - x^1) \leq 0.$$

(iii) Set $\mathcal{O}$ is rationalisable over $\mathbb{R}^\ell_+$ by an increasing, concave function $f : \mathbb{R}^\ell_+ \to \mathbb{R}_+$.

In order to understand how the above condition is related to the supermodular (increasing supermodular) dominance axiom, take any sample $\{(x^s, y^s, p^s)\}_{s\in S}$ of $\mathcal{R}_{X^*}$, where X^* is the smallest (by set inclusion) product lattice containing $\mathcal{X}$. As previously, let distributions ν and $\mu \in \Delta_{X^*}$ be defined by $\nu(z) := \frac{1}{|S|}|\{s \in S : z = x^s\}|$ and $\mu(z) := \frac{1}{|S|}|\{s \in S : z = y^s\}|$. We claim that set $\mathcal{O}$ is cyclically monotone if and only if for any sample $\{(x^s, y^s, p^s)\}_{s\in S}$ of $\mathcal{R}_{X^*}$ such that $\mu = \nu$, we have $\sum_{s\in S} p^s(x^s - y^s) \leq 0$ (denote the latter condition by $*$).

To see that cyclical monotonicity implies ($*$), take any $\{(x^s, y^s, p^s)\}_{s\in S}$ such that $\mu = \nu$. In particular, for all $s \in S$, we have $y^s \in \mathcal{X}$, or equivalently, $(y^s, p) \in \mathcal{O}$, for some p. Clearly, it is possible to construct a sequence $\{(x^n, p^n)\}_{n=1}^m$ of length $m = |S|$

such that (x^1, x^2, p^1), (x^2, x^3, p^2), $\ldots$, (x^m, x^1, p^m) belong to $\mathcal{R}_{X^*}$.[6] Moreover, by cyclical monotonicity, we have

$$p^1 \cdot (x^1 - x^2) + p^2 \cdot (x^2 - x^3) + \cdots + p^m(x^m - x^1) = \sum_{s \in S} p^s(x^s - y^s) \leq 0$$

In order to show that the opposite implication also holds, take any finite sequence $\{(x^n, p^n)\}_{n=1}^m$ in $\mathcal{O}$. Observe that any such sequence generates a sample $\{(x^n, y^n, p^n)\}_{n=1}^m$, where $y^n = x^{n+1}$ for $n = 1, \ldots, m-1$, and $y^m = x^1$. Clearly the corresponding distributions μ and ν are equal. Moreover, condition $(*)$ implies that $\sum_{n=1}^m p^n \cdot (x^n - y^n) \leq 0$.

Given the above characterisation, it is straightforward to show that the supermodular (increasing supermodular) dominance axiom imposes a stronger condition on the set of observations than cyclical monotonicity. Namely, as any two equivalent distributions are at the same time ordered with respect to $\succeq_S$ ($\succeq_{IS}$), in order to verify whether the set of observations obeys the supermodular (increasing supermodular) dominance axiom we need to check the condition stated in the definition of the axiom for a larger class of samples. That is, not only the samples for which the corresponding distributions are equivalent, but also those for which they are ordered with respect to $\succeq_S$ ($\succeq_{IS}$).

3.4.2 Supermodularity and quasilinear rationalisation

The results obtained in the previous sections can be easily reinterpreted as a consumer optimisation problem with quasilinear utility function. Suppose that the consumer's preference can be represented by function $v : \mathbb{R}^\ell_+ \times \mathbb{R}_+ \to \mathbb{R}$, defined by

$$v(x, m) := u(x) + m,$$

where $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ is a utility function defined over ℓ consumption goods, while m denotes the value of the numeraire. Given some initial wealth w, the consumer optimisation problem is

$$\max_{(x,m)} v(x, m), \text{ s.t. } p \cdot x + m \leq w.$$

[6] The sequence can be constructed as follows. Take any element (x, y, p) from $\{(x^s, y^s, p^s)\}_{s \in S}$, and denote it (x^1, x^2, p^1). Given that $\mu = \nu$, there exists a triple (x', y', p') in the sample such that $x' = x^2$. Denote $(x^2, x^3, p^2) := (x', y', p')$. Analogously, we can find some (x'', y'', p'') such that $x'' = x^3$, and so on. Since the sample is finite, eventually we reach the final element (x^m, x^1, p^m).

Clearly, as function v strictly increases with respect to m, the budget constraint is binding at every solution to the above problem. Therefore, without loss of generality, the inequality in the constraint can be replaced with equality.

Suppose that the set of observations is given by $\{(p^k, x^k, m^k)\}_{k\in K}$, where p^k and x^k denote the price and the demand for the consumption goods, while m^k is the chosen level of the numeraire. The set of observations is *rationalisable* by a quasilinear utility function if there is a function $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ such that for any $k \in K$, we have $u(x^k) + m^k \geq u(y) + n$, for all $y \in \mathbb{R}^\ell_+$ and $n \geq 0$ such that $p^k \cdot x^k + m^k = p^k \cdot y + n$. Equivalently, the set is rationalisable in the above sense whenever

$$u(x^k) - p^k \cdot x^k \geq u(y) - p^k \cdot y,$$

for all $y \in \mathbb{R}^\ell_+$ such that $p^k \cdot x^k + m^k \geq p^k \cdot y$. The above objective function very closely resembles the profit function discussed in Section 3.3. In particular, under some additional assumptions, Theorem 3.1 can be applied directly to the above context.

Given the observation set $\{(p^k, x^k, m^k)\}_{k\in K}$, denote the reduced set of observations by $\mathcal{O} := \{(x^k, p^k)\}_{k\in K}$, and let $\mathcal{X}$ be the set of observable demands $\{x^k\}_{k\in K}$. Moreover, let X^* be the smallest (by set inclusion) product lattice containing $\mathcal{X}$, while

$$\mathcal{R}_{X^*} := \{(x, y, p) : (x, p) \in \mathcal{O} \text{ and } y \in X^*\}.$$

Suppose that the set of observations $\{(p^k, x^k, m^k)\}_{k\in K}$ is such that for all $k \in K$, we have $p^k \cdot x^k + m^k \geq p^k \cdot x^l$, for all $l \in K$. Clearly, by applying Proposition 3.6, the data set is rationalisable by a quasilinear utility function if and only if it obeys cyclical monotonicity. Moreover, without loss of generality we may assume that the function is concave and increasing.

In a similar manner we can apply Theorem 3.1 in order to determine the necessary and sufficient conditions for the set of observations to be rationalisable by a supermodular (increasing and supermodular) quasilinear utility function.

Proposition 3.7. *Suppose that the set of observations $\{(p^k, x^k, m^k)\}_{k\in K}$ is such that for all $k \in K$, we have $p^k \cdot x^k + m^k \geq p^k \cdot y$, for all $y \in X^*$. Then, it is rationalisable by a supermodular (increasing and supermodular) quasilinear utility function if and only if set $\mathcal{O}$ obeys the supermodular (increasing supermodular) dominance axiom over X^*.*

Proof. First, we show ($\Rightarrow$). Take any supermodular (increasing and supermodular) function $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ and any sample $\{(x^s, y^s, p^s)\}_{s\in S}$ of $\mathcal{R}_{X^*}$. Since for all $k \in K$, we have $p^k \cdot x^k + m^k \geq p^k \cdot y$, for all $y \in X^*$, it must be that $u(x^s) - p^s \cdot x^s \geq u(y^s) - p^s \cdot y^s$, for all $s \in S$. Define distributions ν and μ over X^* as in the definition of the axiom. Whenever $\mu \succeq_s \nu$ ($\mu \succeq_{IS} \nu$), then $0 \geq \sum_{s\in S} u(x^s) - \sum_{s\in S} u(y^s) \geq \sum_{s\in S} p^s \cdot (x^s - y^s)$.

In order to show ($\Leftarrow$), note that by Theorem 3.1, whenever set $\mathcal{O}$ obeys the supermodular (increasing supermodular) dominance axiom over X^*, there exists a supermodular (increasing supermodular) function $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ such that, for any $(x, p) \in \mathcal{O}$, we have $u(x) - p \cdot x \geq u(y) - p \cdot y$, for all $y \in \mathbb{R}^\ell_+$. In particular, the condition holds for any triple (x^k, p^k, m^k) in the set of observations and any $y \in \mathbb{R}^\ell_+$ such that $p^k \cdot x^k + m^k \geq p^k \cdot y$. □

The assumption stated at the beginning of the proposition is crucial for the result to hold. Namely it requires that every element in set X^* is affordable for the agent at every observation. Equivalently, it implies that the value of the numeraire is high enough for the budget constraint to be neglected while rationalising the set of observations. This allows us to treat the consumer optimisation problem in the exact same way as profit-maximisation discussed in Section 3.3. Once the additional assumption is dropped, the supermodular (increasing supermodular) dominance axiom remains sufficient for the corresponding form of rationalisation, but is no longer necessary. This is due to the fact that, for any observation (x^k, p^k, m^k), condition $u(x^k) - p^k \cdot x^k \geq u(y) - p^k \cdot y$ has to be satisfied only for those $y \in X^*$, for which $p^k \cdot x^k + m^k \geq p^k \cdot y$.

3.5 Concluding remarks

In this chapter, we focused on the testable restrictions for profit-maximisation with production complementarities. Assuming that we observe only a finite number of demands for input factors and their prices, in the main theorem of this part of the thesis we determine a necessary and sufficient condition under which the data set can be rationalised by a supermodular production function. Our axiomatic characterisation of firms' behaviour as well as the easy-to-apply method of verifying the condition provide the foundation for an empirical analysis of "modern manufacturing". Given the importance of the notion of complementarity and its mathematical formalisation for the economic analysis, we consider the our result to be an important step towards verifying or refuting the hypothesis on empirical grounds.

Appendix A

Auxiliary results

In the following chapter we present the auxiliary results that play an important role in our arguments supporting the main claims discussed in the thesis.

We begin by introducing two lemmas employed in the proof of Theorem 1.1 in Chapter 1. In particular, we determine some properties of optimal choices of sophisticated consumers with time-dependent preferences. The notation used in the first two lemmas corresponds to the one presented in Chapter 1.

Lemma A.1. *Take any $t \in T$, $\hat{p}_t \in \hat{P}_t$, and $\hat{x}^i_t$, $\hat{y}^i_t \in \hat{X}_t$ such that, for all $s \geq t$, we have $p_s \cdot y^i_s = p_s \cdot x^i_s$. Then, $F^i_t(\hat{p}_t, \hat{y}^i_t) = F^i_t(\hat{p}_t, \hat{x}^i_t)$ and $V^i_t(\hat{p}_t, \hat{y}^i_t) = V^i_t(\hat{p}_t, \hat{x}^i_t)$.*

Proof. Take any $t \in T$ and any two consumption paths $\hat{x}^i_t$, $\hat{y}^i_t \in \hat{X}_t$ such that, for all $s \geq t$, we have $p_s \cdot y^i_s = p_s \cdot x^i_s$. By definition,

$$F^i_t(\hat{p}_t, \hat{y}^i_t) := \left\{ (z^i_t, \hat{z}^i_{t+1}) \in B_t(\hat{p}_t, \hat{y}^i_t) : \hat{z}^i_{t+1} \in V^i_{t+1}(\hat{p}_{t+1}, \hat{z}^i_{t+1}) \right\}.$$

Clearly, by the initial assumption, we have $B_t(\hat{p}_t, \hat{y}^i_t) = B_t(\hat{p}_t, \hat{x}^i_t)$, which by the above definition implies that $F^i_t(\hat{p}_t, \hat{y}^i_t) = F^i_t(\hat{p}_t, \hat{x}^i_t)$. Since $V^i_t(\hat{p}_t, \hat{y}^i_t)$ and $V^i_t(\hat{p}_t, \hat{x}^i_t)$ contain the $\succeq^i_t$-greatest elements of $F^i_t(\hat{p}_t, \hat{y}^i_t)$ and $F^i_t(\hat{p}_t, \hat{x}^i_t)$ respectively, they are equal. □

Lemma A.2. *Let $\hat{x}^i_t \in V^i_t(\hat{p}_t, \hat{x}^i_t)$. Then, for any $\hat{y}^i_t \in \hat{X}_t$ such that, for all $s \geq t$, $\hat{y}^i_s \sim^i_s \hat{x}^i_s$ and $p_s \cdot y^i_s = p_s \cdot x^i_s$, we have $y^i_t \in V^i_t(\hat{p}_t, \hat{x}^i_t)$.*

Proof. We prove the result by induction. We begin by constructing the base step. Consider the final period $t = T$, and let $x^i_T \in V^i_T(p_T, x^i_T)$. Take any $y^i_T \in X_T$ such that $y^i_T \sim^i_T x^i_T$ and $p_T \cdot y^i_T = p_T \cdot x^i_T$. Clearly, bundle y^i_T belongs to $B_T(p_T, x^i_T)$. Moreover, since $y^i_T \sim^i_T x^i_T$, it must be that $y^i_T \in V^i_T(p_T, x^i_T)$.

Next, take any $t \in T$ and $\hat{x}^i_t \in V^i_t(\hat{p}_t, \hat{x}^i_t)$. Let a consumption path $\hat{y}^i_t \in \hat{X}_t$ be such that, for all $s \geq t$, we have $\hat{y}^i_s \sim^i_s \hat{x}^i_s$ and $p_s \cdot y^i_s = p_s \cdot x^i_s$. In order to construct the inductive step, suppose that $\hat{y}^i_{t+1} \in V^i_t(\hat{p}_{t+1}, \hat{x}^i_{t+1})$. By Lemma A.1, we know that sets $V^i_{t+1}(\hat{p}_{t+1}, \hat{y}^i_{t+1})$ and $V^i_{t+1}(\hat{p}_{t+1}, \hat{x}^i_{t+1})$ are equal. Therefore, it must be that $\hat{y}^i_{t+1} \in V^i_{t+1}(\hat{p}_{t+1}, \hat{y}^i_{t+1})$. Moreover, by definition we have

$$F^i_t(\hat{p}_t, \hat{x}^i_t) := \left\{(z^i_t, \hat{z}^i_{t+1}) \in B_t(\hat{p}_t, \hat{x}^i_t) : \hat{z}^i_{t+1} \in V^i_{t+1}(\hat{p}_{t+1}, \hat{z}^i_{t+1})\right\}.$$

Clearly, $\hat{y}^i_t \in B_t(\hat{p}_t, \hat{x}^i_t)$. Since $\hat{y}^i_{t+1} \in V^i_{t+1}(\hat{p}_{t+1}, \hat{y}^i_{t+1})$, we have that $\hat{y}^i_t \in F^i_t(\hat{p}_t, \hat{x}^i_t)$. Hence, it must be that $\hat{y}^i_t \in V^i_t(\hat{p}_t, \hat{x}^i_t)$, which completes the proof. □

Next, we concentrate on two distinct variations of the well-known Farkas' Lemma. First, we state the so-called *Motzkin's Rational Transposition*, to which we refer in the arguments supporting the results presented in Chapter 2.

Theorem A.1 (Motzkin's Rational Transposition). *Let A be an $k \times n$ rational matrix, B be an $l \times n$ rational matrix, and C be an $m \times n$ rational matrix. Exactly one of the following alternatives is true.*

(i) There exists $x \in \mathbb{R}^n$ such that $A \cdot x \gg 0$, $B \cdot x \geq 0$, and $C \cdot x = 0$.

(ii) There exist $\theta \in \mathbb{Z}^k$, $\lambda \in \mathbb{Z}^l$, and $\pi \in \mathbb{Z}^m$ such that $\theta \cdot A + \lambda \cdot B + \pi \cdot C = 0$, where $\theta > 0$ and $\lambda \geq 0$.

The proof of the above theorem can be found in Stoer and Witzgall (1970). Our next result is a specific version of Farkas' Lemma, which plays a substantial role in the proof of Proposition 3.4 and Theorem 3.1 stated in Chapter 3.

Theorem A.2 (Farkas' Lemma). *Let A be a $m \times n$ rational matrix and let $b \in \mathbb{R}^m$. Exactly one of the following alternatives is true.*

(i) There exists $x \in \mathbb{R}^n$ such that $A \cdot x \geq b$.

(ii) There exists $\lambda \in \mathbb{Q}^m$, where $\lambda > 0$, such that $\lambda \cdot A = 0$ and $\lambda \cdot b > 0$.

The above result is a special case of the rational version of Farkas' Lemma. However, in the above theorem we allow for vector b to be an element of $\mathbb{R}^m$ rather than $\mathbb{Q}^m$. This variation on the original result is significant for the main argument of Chapter 3. Clearly, without loss of generality, we can substitute $\mathbb{Q}^m$ with $\mathbb{Z}^m$. We prove the result using Fourier-Motzkin's Elimination Theorem. In order to make our thesis self-contained, we present a modified version of the theorem below. The original proof can be found in Fourier (1826) or Motzkin (1951).

Lemma A.3 (Fourier-Motzkin's Elimination). *Define* $P := \{x \in \mathbb{R}^n : A \cdot x \geq b\}$, *where* A *is a* $m \times n$ *rational matrix and* $b \in \mathbb{R}^m$, *and let* $\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ *be a projection defined by* $\pi(x^1, x^2, \ldots, x^n) := (x^2, \ldots, x^n)$. *There exists a positive rational matrix* Λ *such that*

$$\pi(P) = \{(x_2, \ldots, x_n) \in \mathbb{R}^{n-1} : \Lambda \cdot A \cdot (x_1, x_2, \ldots, x_n) \geq \Lambda \cdot b, \ \textit{for all } x_1 \in \mathbb{R}\}.$$

Proof. Denote the entries of matrix A by a_{ij}. By definition, set P consists of all the elements x in $\mathbb{R}^n$ that satisfy

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \geq b_i, \text{ for all } i = 1, \ldots, m.$$

Define sets $G := \{i : a_{i1} > 0\}$, $Z := \{i : a_{i1} = 0\}$, and $L := \{i : a_{i1} < 0\}$, and reformulate every inequality $k \in G$ such that

$$x_1 \geq -\frac{a_{k2}}{|a_{k1}|}x_2 - \cdots - \frac{a_{kn}}{|a_{k1}|}x_n - \frac{b_k}{|a_{k1}|}.$$

Moreover, reformulate any inequality in $l \in L$ so that

$$x_1 \leq \frac{a_{l2}}{|a_{l1}|}x_2 + \cdots + \frac{a_{ln}}{|a_{l1}|}x_n + \frac{b_l}{|a_{l1}|},$$

The above reformulation allows us to reduce $|G| + |L|$ inequalities corresponding to the indices in sets G and L to just two inequalities defined by

$$\min_{l \in L} \left\{ \frac{a_{l2}}{|a_{l1}|}x_2 + \cdots + \frac{a_{ln}}{|a_{l1}|}x_n + \frac{b_l}{|a_{l1}|} \right\} \geq x_1 \geq \max_{k \in G} \left\{ -\frac{a_{k2}}{|a_{k1}|}x_2 - \cdots - \frac{a_{kn}}{|a_{k1}|}x_n - \frac{b_k}{|a_{k1}|} \right\}.$$

Observe that vector $(x_2, \ldots, x_n)$ belongs to projection $\pi(P)$ if and only if there exists some real number x_1 such that $(x_1, x_2, \ldots, x_n) \in P$. This implies that vector $(x_2, \ldots, x_n)$ has to satisfy all the inequalities in Z as well as

$$\min_{l \in L} \left\{ \frac{a_{l2}}{|a_{l1}|}x_2 + \cdots + \frac{a_{ln}}{|a_{l1}|}x_n + \frac{b_l}{|a_{l1}|} \right\} \geq \max_{k \in G} \left\{ -\frac{a_{k2}}{|a_{k1}|}x_2 - \cdots - \frac{a_{kn}}{|a_{k1}|}x_n - \frac{b_k}{|a_{k1}|} \right\}.$$

Moreover, vector $(x_2, \ldots, x_n)$ satisfies the above inequality if and only if it satisfies $|G||L|$ inequalities such that for any $k \in G$ and $l \in L$, we have

$$\frac{a_{l2}}{|a_{l1}|}x_2 + \cdots + \frac{a_{ln}}{|a_{l1}|}x_n + \frac{b_l}{|a_{l1}|} \geq -\frac{a_{k2}}{|a_{k1}|}x_2 - \cdots - \frac{a_{kn}}{|a_{k1}|}x_n - \frac{b_k}{|a_{k1}|}$$

or equivalently

$$\left(\frac{a_{k2}}{|a_{k1}|}+\frac{a_{l2}}{|a_{l1}|}\right)x_2+\cdots+\left(\frac{a_{kn}}{|a_{k1}|}+\frac{a_{ln}}{|a_{l1}|}\right)x_n\geq\left(\frac{b_k}{|a_{k1}|}+\frac{b_l}{|a_{l1}|}\right).$$

Define a $|G||L|+|Z|$ by m matrix Λ as follows. For each of the first $|G||L|$ rows, take the corresponding indices $k\in G$ and $l\in L$. Let the k'th entry of that row be equal to $1/|a_{k1}|$, the l'th entry be equal to $1/|a_{l1}|$, and all the remaining entries be equal to zero. Finally, in each of the $|Z|$ remaining rows, let the entry corresponding to element $z\in Z$ be equal to 1, and all the remaining entries be equal to zero.

Clearly, Λ is positive. In addition, since A is rational, so is Λ. In order to conclude the proof, note that, by construction, vector $(x_2,\ldots,x_n)$ solves the above system of inequalities if and only if $(\Lambda\cdot A)\cdot(x_1,x_2,\ldots,x_n)\geq(\Lambda\cdot b)$, for all $x_1\in\mathbb{R}$. Hence, $(x_2,\ldots,x_n)\in\pi(P)$ if and only if $(\Lambda\cdot A)\cdot(x_1,x_2,\ldots,x_n)\geq(\Lambda\cdot b)$, for any $x_1\in\mathbb{R}$. □

We proceed with the proof of Theorem A.2.

Proof of Theorem A.2. We begin by showing (i). Suppose that there is no $\lambda\in\mathbb{Q}^m$, $\lambda>0$, such that $\lambda\cdot A=0$ and $\lambda\cdot b>0$. Define set $P:=\{x\in\mathbb{R}^n:A\cdot x\geq b\}$, where A and b are specified as in the thesis of the theorem. Hence, $A\cdot x\geq b$, for some $x\in\mathbb{R}^n$ if and only if $x\in P$. Let $\pi^\ell:\mathbb{R}^\ell\to\mathbb{R}^{\ell-1}$ be a projection defined by $\pi^\ell(x_1,x_2,\ldots,x_\ell)=(x_2,\ldots,x_\ell)$, for some $\ell>1$. Clearly, set P is non-empty if and only if set $\pi^2\circ\pi^3\circ\cdots\circ\pi^n(P)\subseteq\mathbb{R}$ is non-empty.

By Lemma A.3, there is a sequence of positive rational matrices $\{\Lambda_i\}_{i=2}^n$ such that

$$\begin{aligned}\pi^2\circ\pi^3\circ\cdots\circ\pi^n(P)={}&\\&\big\{x_n\in\mathbb{R}:(\Lambda_2\cdots\Lambda_n\cdot A)\cdot(x_1,\ldots,x_n)\geq(\Lambda_2\cdots\Lambda_n\cdot b),\\&\qquad\text{for all }(x_1,\ldots,x_{n-1})\in\mathbb{R}^{n-1}\big\}.\end{aligned}$$

Denote $C:=(\Lambda_2\cdots\Lambda_n\cdot A)$ and $d=(\Lambda_2\cdots\Lambda_n\cdot b)$. Clearly, matrix C has n columns. Denote the number of rows of the matrix and the dimension of vector d by r. Note that matrix C is rational, with non-zero entries appearing only in the n'th column. We shall denote the entries of the matrix by c_{ij}, and the entries of vector d by d_i, where $i=1,\ldots,r$ and $j=1,\ldots,n$.

Define sets $G:=\{i:c_{in}>0\}$, $Z:=\{i:c_{in}=0\}$, and $L:=\{i:c_{in}<0\}$. Observe that, for any $i\in Z$, it must be that $d_i\leq 0$. Otherwise, there would exist a row of matrix $(\Lambda_2 \cdots \Lambda_n)$, denoted by λ, such that $\lambda \cdot A = 0$ and $\lambda \cdot b > 0$. Since λ is positive and rational, this would lead to a contradiction.

Set $\pi^2 \circ \pi^3 \circ \cdots \circ \pi^n(P)$ is non-empty if and only if $\min_{l \in L} d_l/c_{ln} \geq \max_{k \in G} d_k/c_{kn}$. Suppose that $\min_{l \in L} d_l/c_{ln} < \max_{k \in G} d_k/c_{kn}$. Take any $l \in L$ and $k \in K$ such that $d_l/c_{ln} < d_k/c_{kn}$ and construct a vector $\gamma \in \mathbb{Q}^r$ with the l'th and k'th entry equal to $1/|c_{ln}|$ and $1/|c_{kn}|$ respectively, while every other entry is equal to zero. Clearly, $\lambda := \gamma \cdot \Lambda_2 \cdots \Lambda_n \in \mathbb{Q}^m$ is a positive vector. Moreover, by definition, $\lambda \cdot A = 0$ while $\lambda \cdot b > 0$. Contradiction. Therefore, whenever there is no $\lambda \in \mathbb{Q}^m$, $\lambda > 0$, such that $\lambda \cdot A = 0$ and $\lambda \cdot b > 0$, there exists some $x \in \mathbb{R}^n_+$ such that $A \cdot x \geq b$.

In order to prove the second part of the result, suppose that there is some $\lambda \in \mathbb{Q}^m$, $\lambda > 0$, such that $\lambda \cdot A = 0$ and $\lambda \cdot b > 0$. Whenever there exists some $x \in \mathbb{R}^n$ such that $A \cdot x \geq b$, we have $(\lambda \cdot b) > (\lambda \cdot A) \cdot x \geq (\lambda \cdot b)$, which yields a contradiction. $\square$

Appendix B

Proofs

In the following section we present proofs of the results stated in Chapter 2 that were not included in the main body of the thesis.

B.1 Proof of Proposition 2.1

In this section we concentrate on the argument supporting Proposition 2.1. Since the necessity of cyclical consistency for rationalisation of the set of observations was already discussed in the main body of the thesis, we concentrate on the sufficiency part of the proposition. We present the argument via three lemmas that follow. In the first result we show that cyclical consistency axiom implies the existence of a consistent, mixed-monotone pre-order on $\mathcal{O}$.

Lemma B.1. *Whenever the set of observations $\mathcal{O}$ is cyclically consistent, it admits a consistent, mixed-monotone pre-order $\mathcal{R}$ over $\mathcal{A}$.*

Proof. First, we define an equivalence relation on $\mathcal{A}$. We will say that elements x and y are related to each other whenever $x\,\mathcal{I}^*y$. Denote a typical equivalence class by $[x]$. We prove our claim by induction on the number of equivalence classes. If there is a single equivalence class of $\mathcal{A}$, then let $x\mathcal{R}y$ for any two elements in $\mathcal{A}$, which is consistent with the revealed preference relation and mixed-monotone. Otherwise, there would be some x, y in $\mathcal{A}$ such that $x\,\mathcal{R}^*y$, $y\,\mathcal{R}^*x$, and $x >_X y$, which would violate cyclical consistency.

Suppose that the claim holds for L equivalence classes of $\mathcal{A}$. We will show that it also holds for $L+1$ classes. Define a relation $\succeq$ over the equivalence classes in the following way. We say that $[x]' \succeq [x]$ whenever there exists some $x' \in [x]'$, $x \in [x]$, and a sequence $\{x^i\}_{i=1}^n$ in $\mathcal{X}$ such that: (i) $x^{i+1}\mathcal{R}^*x^i$ or $x^{i+1} \geq_X x^i$, (ii) $x'\mathcal{R}^*x^1$ or $x' \geq_X x^1$, and (iii) $x^n\mathcal{R}^*x$ or $x^n \geq_X x$. Clearly, the relation is transitive. In order to show that it is anti-symmetric, assume the opposite, i.e., that there exist two distinct equivalence classes $[x]'$ and $[x]$ such that $[x]' \succeq [x]$ and $[x] \succeq [x]'$. The first relation implies that there exist some $x' \in [x]'$, $x \in [x]$, and a sequence $\{x^i\}_{i=1}^n$ satisfying conditions (i), (ii), and (iii). The second relation implies that there exist some $y \in [x]$, $y' \in [x]'$, and a sequence $\{y^i\}_{i=1}^m$ satisfying conditions (i), (ii), and (iii). Hence, there exists a sequence $\{z^i\}_{i=1}^k$ in $\mathcal{A}$ such that every subsequent element of the sequence dominates the preceding one with respect to $\mathcal{R}^*$ or $\geq_X$, and $z^1 \geq_X z^k$. If the elements are ordered only with respect to $\mathcal{R}^*$, then it must be $[x]' = [x]$. This yields a contradiction, since by assumption the two equivalence classes are distinct. Otherwise, we have $z^1 >_X z^k$, which violates cyclical consistency.

Given the above argument, it follows that there exists an equivalence class $\underline{[x]}$ such that there is no other equivalence class $[x]$ with $\underline{[x]} \succeq [x]$. Moreover, the exclusion of the equivalence class from $\mathcal{A}$ does not affect the relations between the remaining equivalence classes. Construct the new set of observations $\mathcal{O}'$ in the following way. Whenever $(A, x) \in \mathcal{O}$ and $x \in \underline{[x]}$, we have $(A, x) \notin \mathcal{O}'$. For any $(A, x) \in \mathcal{O}$ such that $x \notin \underline{[x]}$, define set $A' := \{x \in A : x \notin \underline{[x]}\}$, and let $(A', x) \in \mathcal{O}'$. Note that the set of observable choices corresponding to $\mathcal{O}'$ is $\mathcal{A}' = \mathcal{A}\backslash\underline{[x]}$. Since $\mathcal{A}'$ has only K equivalence classes, by the induction hypothesis we conclude that there exists a mixed-monotone pre-order $\mathcal{R}$ on $\mathcal{A}'$ consistent with the revealed relations generated by $\mathcal{O}'$. We can extend the pre-order to the whole set $\mathcal{A}$ by letting $x'\mathcal{P}x$ for all $x' \in \mathcal{A}'$ and $x \in \underline{[x]}$, and $x'\mathcal{I}x$, for x', $x \in \underline{[x]}$. □

In the following lemma we show that whenever the set of observation admits a mixed-monotone pre-order over $\mathcal{A}$, there exists a solution to a certain system of linear inequalities. Recall that $\bar{\mathcal{A}} := \mathcal{M} \times \mathcal{T}$.

Lemma B.2. *Let $\mathcal{O}$ admit a consistent, mixed-monotone pre-order $\mathcal{R}$ on $\mathcal{A}$. There are numbers $\{v_m^t\}_{(m,t)\in\bar{\mathcal{A}}}$ such that for any (m, t), $(n, s) \in \bar{\mathcal{A}}$, (i) if $(m, t) >_X (n, s)$ or $(m, t)\mathcal{P}(n, s)$ then $v_m^t > v_n^s$, and (ii) $(m, t)\mathcal{I}(n, s)$ implies $v_m^t = v_n^s$.*

Proof. Since $\mathcal{O}$ admits a mixed-monotone pre-order $\mathcal{R}$ consistent with the revealed relations, for any two elements x, y in $\mathcal{A}$, we have either $x\mathcal{P}y$ or $x\mathcal{I}y$. First, we determine numbers $\{v_m^t\}_{(m,t)\in\mathcal{A}}$ in a recursive manner. Take any $(n,s)\in\mathcal{A}$ such that for all $(m,t)\in\mathcal{A}$ we have $(m,t)\mathcal{R}(n,s)$. Clearly, such element exists. Assign any strictly positive value to v_n^s and define $\mathcal{X}_1 := \{(n,s)\}$ as well as $V_1 := \{v_n^s, 0\}$.

For any $j \geq 1$, assume that $\mathcal{X}_j \subseteq \mathcal{A}$ is non-empty, and for all $(m,t)\in\mathcal{A}\backslash\mathcal{X}_j$, we have $(m,t)\mathcal{R}(n,s)$, for any $(n,s)\in\mathcal{X}_j$. Moreover, assume that set V_j is a finite set of strictly positive real numbers and 0. Take any $(n,s)\in\mathcal{A}\backslash\mathcal{X}_j$ such that for all $(m,t)\in\mathcal{A}\backslash\mathcal{X}_j$, we have $(m,t)\mathcal{R}(n,s)$. If there exists a $(m,t)\in\mathcal{X}_j$ such that $(n,s)\mathcal{I}(m,t)$, let $v_n^s = \max V_j$. Otherwise, set v_n^s to be any number strictly greater than $\max V_j$. Denote $\mathcal{X}_{j+1} := \mathcal{X}_j \cup \{(n,s)\}$ and $V_{j+1} := V_j \cup \{v_n^s\}$. The algorithm terminates whenever $\mathcal{A} = \mathcal{X}_j$. In this case, denote $V := V_j$. By construction, for any (m,t), $(n,s)\in\mathcal{A}$, if $(m,t)\mathcal{P}(n,s)$ then $v_m^t > v_n^s$, and $(m,t)\mathcal{I}(n,s)$ implies $v_m^t = v_n^s$.

In order to complete the proof, we need to determine the values of the remaining elements of $\{v_m^t\}_{(m,t)\in\bar{\mathcal{A}}}$. If $(\overline{m},\underline{t})$ belongs to $\mathcal{A}$, let $v_{\overline{m}}^{\underline{t}}$ take the value assigned previously. Otherwise, let $v_{\overline{m}}^{\underline{t}}$ be any number strictly greater than $\max V$. We determine the remaining numbers in the following way. If $(m,t)\in\mathcal{A}$, then v_m^t takes the value as determined above. Otherwise, define $m^+ := \min\{n\in\mathcal{M} : n > m\}$ and $t^- := \max\{s\in\mathcal{T} : s < t\}$.[1] Set $v_m^t = v$, where v is any number satisfying

$$\min\left\{v_{m^+}^t, v_m^{t^-}\right\} > v > \max\left\{v'\in V : \min\left\{v_{m^+}^t, v_m^{t^-}\right\} > v'\right\}.$$

Finally, we show that, for any (m,t), both $v_m^t < v_{m^+}^t$ and $v_m^t < v_m^{t^-}$. If (m,t) does not belong to $\mathcal{A}$, the claim holds by construction. Therefore, it suffices to consider the case when $(m,t)\in\mathcal{A}$. We prove the claim by contradiction. Suppose that $v_m^t \geq v_{m^+}^t$ or $v_m^t \geq v_m^{t^-}$. This would imply, that there exists some $n \geq m$ and $s \leq t$ such that $(n,s)\in\mathcal{A}$ and $v_n^s \leq v_m^t$. By construction of $\{v_m^t\}_{(m,t)\in\mathcal{A}}$, this would mean that $(m,t)\mathcal{R}(n,s)$. However, unless $(m,t) = (n,s)$, we have $(n,s) >_X (m,t)$. Since $\mathcal{R}$ is mixed-monotone, this yields a contradiction. □

Clearly, any sequence that satisfies the properties discussed in the previous lemma automatically determines a utility function defined over $\bar{\mathcal{A}}$ that rationalises the set of observations. In the following lemma we show how to extend the function to X.

[1] That is, m^+ is the immediate successor of m, while t^- is the immediate predecessor of t (with respect to the increasing order on $\mathcal{M}$ and $\mathcal{T}$ respectively).

Lemma B.3. *Set $\mathcal{O}$ is rational if there exists a sequence of numbers $\{v_m^t\}_{(m,t)\in\bar{\mathcal{A}}}$ such that, for any (m,t), $(n,s) \in \bar{\mathcal{A}}$, whenever $(m,t) >_X (n,s)$ then $v_m^t > v_n^s$, and $(m,t)\,\mathcal{R}^*(n,s)$ implies $v_m^t \geq v_n^s$.*

Proof. Whenever $\underline{m} \neq 0$, construct a decreasing sequence $\{v_0^t\}_{t\in\mathcal{T}}$ such that, for all $t \in \mathcal{T}$, we have $v_0^t < v_{\underline{m}}^t$. Similarly, whenever $\underline{t} \neq 0$, let $\{v_m^0\}_{m\in\mathcal{M}}$ be a decreasing sequence with $v_m^0 > v_m^{\underline{t}}$, for all $m \in \mathcal{M}$. If both $\underline{m} \neq 0$ and $\underline{t} \neq 0$, let v_0^0 be such that $v_0^{\underline{t}} < v_0^0 < v_{\underline{m}}^0$. Finally, denote $\mathcal{M}_0 := \mathcal{M} \cup \{0\}$ and $\mathcal{T}_0 := \mathcal{T} \cup \{0\}$.

For any $m \in \mathcal{M}_0 \backslash \{\overline{m}\}$, define $m^+ := \min\{n \in \mathcal{M} : n > m\}$. Similarly, for any $t \in \mathcal{T}_0 \backslash \{\bar{t}\}$, let $t^+ := \min\{s \in \mathcal{T} : s > t\}$. We begin by defining function $w : [0,\overline{m}] \times [0,\bar{t}] \to \mathbb{R}$. For any $m \in \mathcal{M}_0 \backslash \{\overline{m}\}$ and $t \in \mathcal{T}_0 \backslash \{\bar{t}\}$, take any (n,s) in $[m,m^+] \times [t,t^+]$ and let

$$w(n,s) := \tfrac{n-m}{m^+-m} \cdot \tfrac{s-t}{t^+-t} \cdot v_m^t + \left(1 - \tfrac{n-m}{m^+-m}\right) \cdot \tfrac{s-t}{t^+-t} \cdot v_{m^+}^t + \tfrac{n-m}{m^+-m} \cdot \left(1 - \tfrac{s-t}{t^+-t}\right) \cdot v_m^{t^+} + \left(1 - \tfrac{n-m}{m^+-m}\right) \cdot \left(1 - \tfrac{s-t}{t^+-t}\right) \cdot v_{m^+}^t.$$

Clearly, the function is defined by a bi-linear combination of the elements that belong to the sequence $\{v_m^t\}_{\mathcal{M}_0\times\mathcal{T}_0}$. In particular, the function is continuous and monotone with respect to $\geq_X$.

Let $\bar{w} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ be an extension of function w to $\mathbb{R}_+ \times \mathbb{R}_+$, defined as follows. For any $(n,s) \in [0,\overline{m}] \times [0,\bar{t}]$, let $\bar{w}(n,s) := w(n,s)$. Let $\{\lambda^t\}_{t\in\mathcal{T}_0}$ be any strictly increasing sequence of strictly positive real numbers, while $\{\mu_m\}_{m\in\mathcal{M}_0}$ denotes any strictly decreasing sequence of strictly negative real numbers. For any $t \in \mathcal{T}_0 \backslash \{\bar{t}\}$, and any $s \in [t,t^+]$, $n \geq \overline{m}$, define

$$\bar{w}(n,s) := \tfrac{s-t}{t^+-t} \cdot \left(v_{\overline{m}}^t + \lambda^t(n - \overline{m})\right) + \left(1 - \tfrac{s-t}{t^+-t}\right) \cdot \left(v_{\overline{m}}^{t^+} + \lambda^{t^+}(n - \overline{m})\right).$$

Analogously, for any $m \in \mathcal{M}_0 \backslash \{\overline{m}\}$, and any $n \in [m,m^+]$, $t \geq \bar{t}$, define

$$\bar{w}(n,s) := \tfrac{n-m}{m^+-m} \cdot \left(v_m^{\bar{t}} + \mu_m(s - \bar{t})\right) + \left(1 - \tfrac{n-m}{m^+-m}\right) \cdot \left(v_{m^+}^{\bar{t}} + \mu_{m^+}(s - \bar{t})\right).$$

Finally, for any $n \geq \overline{m}$ and $s \geq \bar{t}$, let $\bar{w}(n,s) := v_{\overline{m}}^{\bar{t}} + \lambda^{\bar{t}}(n - \overline{m}) + \mu_{\overline{m}}(s - \bar{t})$. Clearly, the above function is continuous and strictly increasing with respect to the partial order $\geq_X$. Moreover, for any $(m,t) \in \bar{\mathcal{A}}$, we have $\bar{w} = v_m^t$. In particular, whenever $x\,\mathcal{R}^*y$ then $\bar{w}(x) \geq \bar{w}(y)$. In order to complete the proof, let $v : \mathbb{R}_+ \times \mathbb{N} \to \mathbb{R}$ be defined by $v(m,t) := \bar{w}(m,t)$, for any point (m,t) in its domain. Obviously, function

v is mixed-monotone and for any $(m,t) \in \bar{\mathcal{A}}$, we have $v(m,t) = v_m^t$. Hence, $x\,\mathcal{R}^* y$ implies $v(x) \geq v(y)$. The proof is complete. □

Note that Lemmas B.1, B.2, and B.3 imply that whenever a set of observations is cyclically consistent it is possible to construct a mixed-monotone utility function $v : X \to \mathbb{R}$ that rationalises $\mathcal{O}$. Hence, cyclical consistency is a sufficient condition for this form of rationalisation. Moreover, note that Lemma B.2 allows to verify the cyclical consistency axiom by finding a solution to a system of linear inequalities. This plays an important role for the applicability of our result.

B.2 Proof of Theorem 2.1

In this section we prove the sufficiency part of Theorem 2.1. That is, we show that the dominance axiom is a sufficient condition for the set of observations to be rationalisable by a discounted utility function. We start with the following lemma.

Lemma B.4. *Whenever $\mathcal{O}$ obeys the dominance axiom there is a strictly increasing sequence $\{\phi_m\}_{m\in\mathcal{M}}$ and a strictly decreasing sequence $\{\varphi_t\}_{t\in\mathcal{T}}$ of real numbers such that $(m,t)\,\mathcal{R}^*(n,s)$ implies $\phi_m + \varphi_t \geq \phi_n + \varphi_s$.*

Proof. Enumerate the elements of $\mathcal{R}^*$ so that it is equal to $\{(x^j, y^j)\}_{j\in J}$, where we denote $x^j = (m^j, t^j)$ and $y^j = (n^j, s^j)$. Let $\mu_m \in \{0,1\}^{|\mathcal{M}|}$, $m \in \mathcal{M}$, be a vector equal to 1 at the coordinate corresponding to m and 0 elsewhere. Analogously define $\tau_t \in \{0,1\}^{|\mathcal{T}|}$, $t \in \mathcal{T}$.

Let $\mathbb{I}$ be a $|\mathcal{M}| + |\mathcal{T}|$ by $|\mathcal{M}| + |\mathcal{T}|$ identity matrix. Moreover, let B denote a $|J|$ times $|\mathcal{M}| + |\mathcal{T}|$ matrix such that, for any $j \in J$, the j'th row of the matrix is equal to $\left(\sum_{k\leq m^j} \mu_k - \sum_{k\leq n^j} \mu_k, \sum_{k\geq t^j} \tau_k - \sum_{k\geq s^j} \tau_k\right)$. We claim that if $\mathcal{O}$ obeys the dominance axiom, there are vectors $\xi \in \mathbb{R}^{|\mathcal{M}|}$ and $\vartheta \in \mathbb{R}^{|\mathcal{T}|}$ such that $\mathbb{I} \cdot (\xi, \vartheta) \gg 0$ and $B \cdot (\xi, \vartheta) \geq 0$.

Suppose that $\mathcal{O}$ obeys the axiom, but there are no such vectors. By Theorem A.1 (see Appendix A), there are some $\theta \in \mathbb{Z}^{|\mathcal{M}|+|\mathcal{T}|}$ and $\lambda \in \mathbb{Z}^{|J|}$ such that $\theta\cdot\mathbb{I} + \lambda\cdot B = 0$ $(\star)$, with $\theta > 0$ and $\lambda \geq 0$. Take any such θ and λ, and let $\lambda = (\lambda_j)_{j\in J}$.

For all $j \in J$, take λ_j copies of pair $(x^j, y^j) \in \mathcal{R}^*$ and construct a sample $\{(x^i, y^i)\}_{i\in I}$, where $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$. Since $\theta > 0$, for equation $(\star)$ to hold it must be that $\sum_{i\in I}\sum_{k\leq m^i}\mu_k \leq \sum_{i\in I}\sum_{k\leq n^i}\mu_k$ and $\sum_{i\in I}\sum_{k\geq t^i}\tau_k \leq \sum_{i\in I}\sum_{k\geq s^i}\tau_k$ is satisfied, with at least one inequality being strict. This is possible only if

$$|\{i\in I : m^i \geq m\}| \leq |\{i\in I : n^i \geq m\}| \quad\text{and}\quad |\{i\in I : t^i \leq t\}| \leq |\{i\in I : s^i \leq t\}|,$$

for all $m \in \mathcal{M}$ and $t \in \mathcal{T}$, while for at least one m or t the inequality is strict. However, this violates the dominance axiom.

To complete the argument, take any vectors ξ and ϑ such that $\mathbb{I}\cdot(\xi,\vartheta) \gg 0$ and $B\cdot(\xi,\vartheta)\geq 0$. Define sequences $\{\xi_m\}_{m\in\mathcal{M}}$ and $\{\vartheta_t\}_{t\in\mathcal{T}}$ by $\xi_m := \mu_m\cdot\xi$ and $\vartheta_t := \tau_t\cdot\vartheta$. By construction, $(m,t)\,\mathcal{R}^*(n,s)$ implies $\sum_{k\leq m}\xi_k + \sum_{k\geq t}\vartheta_k \geq \sum_{k\leq n}\xi_k + \sum_{k\geq s}\vartheta_k$. Let $\phi_m := \sum_{k\leq m}\xi_k$ and $\varphi_t := \sum_{k\geq t}\vartheta_k$, for all $m\in\mathcal{M}$ and $t\in\mathcal{T}$. Since sequences $\{\xi_m\}_{m\in\mathcal{M}}$ and $\{\vartheta_t\}_{t\in\mathcal{T}}$ are strictly positive, both $\{\phi_m\}_{m\in\mathcal{M}}$ and $\{\varphi_t\}_{t\in\mathcal{T}}$ are strictly monotone. Moreover, $(m,t)\,\mathcal{R}^*(n,s)$ implies $\phi_m+\varphi_t \geq \phi_n+\varphi_s$. □

In the next lemma we show how to construct a discounted utility function that rationalises the set of observations

Lemma B.5. *Set $\mathcal{O}$ is rationalisable by a discounted utility function whenever there is a strictly increasing sequence $\{\phi_m\}_{m\in\mathcal{M}}$ and a strictly decreasing sequence $\{\varphi_t\}_{t\in\mathcal{T}}$ of real numbers such that $(m,t)\,\mathcal{R}^*(n,s)$ implies $\phi_m+\varphi_t\geq\phi_n+\varphi_s$.*

Proof. Take any number $\phi_0 < \phi_{\underline{m}}$ whenever $\underline{m}\neq 0$, and let $\phi_0=\phi_{\underline{m}}$ otherwise. Denote $\mathcal{M}_0=\mathcal{M}\cup\{0\}$. In addition, for any $m\in\mathcal{M}\cup\{0\}$ different from $\overline{m}$, define the immediate successor of m by $m^+ := \min\{n\in\mathcal{M} : n>m\}$. Let $\phi:\mathbb{R}\to\mathbb{R}$ be specified by

$$\phi(n) := \sum_{m\in\mathcal{M}_0}[\phi_m + \lambda_m(n-m)]\chi_{N_m}(n), \tag{B.1}$$

where χ is the indicator function,[2] $\lambda_m := (\phi_{m^+}-\phi_m)/(m^+-m)$ for all m different from $\overline{m}$, $\lambda_{\overline{m}}$ is any strictly positive number, while $N_m := [m,m^+)$ for $m\neq\overline{m}$, and $N_{\overline{m}} := [\overline{m},\infty)$. Clearly, the function is continuous and strictly increasing. Moreover, for any $m\in\mathcal{M}$, we have $\phi(m)=\phi_m$.

If $\underline{t}=0$, define sequence $\{\tilde{\varphi}_t\}_{t\in\mathcal{T}}$ by $\tilde{\varphi}_t := \varphi_t-\varphi_{\underline{t}}$. Otherwise, let $\tilde{\varphi}_t := \varphi_t$. Denote $\tilde{\varphi}_0=0$, $\mathcal{T}_0 := \mathcal{T}\cup\{0\}$, and let the immediate successor of t in $\mathcal{T}$ be defined

[2]That is, for any set Y, we have $\chi_Y(x)=1$ whenever $x\in Y$, and $\chi_Y(x)=0$ otherwise.

by $t^+ := \min\{s \in \mathcal{T} : s > t\}$, for any $t \neq \bar{t}$. Define function $\varphi : \mathbb{N} \to \mathbb{R}_-$ by

$$\varphi(s) := \sum_{t \in \mathcal{T} \cup \{0\}} [\tilde{\varphi}_t + v_t(s-t)]\chi_{S_t}(s)$$

where χ is the indicator function, $v_t := (\tilde{\varphi}_{t^+} - \tilde{\varphi}_t)/(t^+ - t)$ for all $t \in \mathcal{T} \cup \{0\}$ different from $\bar{t}$, $v_{\bar{t}}$ is any strictly negative number, while $S_t := \{s \in \mathbb{N} : t \leq s < t^+\}$ for $t \neq \bar{t}$, and $S_{\bar{t}} := \{s \in \mathbb{N} : \bar{t} \leq s\}$. Note that the function is strictly decreasing and takes only negative values with $\varphi(0) = 0$. Moreover, we have $\varphi(t) = \tilde{\varphi}_t$, for any $t \in \mathcal{T}$.

Define functions $u : \mathbb{R}_+ \to \mathbb{R}_+$ and $\gamma : \mathbb{N} \to (0,1]$ by $u := \exp(\phi)$ and $\gamma := \exp(\varphi)$. Clearly the two functions are strictly monotone, while $\gamma(0) = 1$. Finally, for any element (m,t) in $\mathcal{A}$, we have $v(m,t) = \exp(\phi_m + \tilde{\varphi}_t)$. Hence, whenever we have $(m,t)\,\mathcal{R}^*(n,s)$ then $v(m,t) \geq v(n,s)$. □

Lemmas B.4 and B.5 imply that whenever the set of observations obeys the dominance axiom, we can construct an instantaneous utility function $u : \mathbb{R}_+ \to \mathbb{R}_+$ and some discounting function $\gamma : \mathbb{N} \to (0,1]$ such that $v(m,t) := u(m)\gamma(t)$ rationalises $\mathcal{O}$. Clearly, this completes the argument supporting Theorem 2.1.

B.3 Proof of Proposition 2.3

We devote this section to the proof of Proposition 2.3. The following result states an important property of anchored experiments. We apply the following lemma in the proof of Proposition 2.3, which is presented in the remainder of this section.

Lemma B.6. *Let $\mathcal{E}$ be an anchored experiment. Whenever $\mathcal{O}$ is rationalisable by a discounted utility function, there is a strictly increasing sequence $\{n_t\}_{t \in \mathcal{T}}$ in $\mathbb{R}_+$ such that $(m,t)\,\mathcal{R}^*(m^*,t^*)$ implies $m \geq n_t$, and $(m^*,t^*)\,\mathcal{R}^*(m,t)$ implies $n_t \geq m$.*

Proof. As set $\mathcal{O}$ is rationalisable by a discounted utility function, there is some strictly increasing $u : \mathbb{R}_+ \to \mathbb{R}_+$ and a strictly decreasing $\gamma : \mathbb{N} \to (0,1]$, with $\gamma(0) = 1$, such that $(m,t)\,\mathcal{R}^*(m^*,t^*)$ implies $u(m)\gamma(t) \geq u(m^*)\gamma(t^*)$, and $(m^*,t^*)\,\mathcal{R}^*(m,t)$ implies $u(m^*)\gamma(t^*) \geq u(m)\gamma(t)$. Let $u^* := u(m^*)$ and $\gamma^* := \gamma(t^*)$, while for all $t \in \mathcal{T}$,

$$n_t = \min\{m \in \mathbb{R}_+ : u(m) \geq u^*\gamma^*/\gamma(t)\}.$$

Since u is continuous and strictly increasing, number n_t is well defined with $u(n_t) = u^*\gamma^*/\gamma(t)$, for all $t \in \mathcal{T}$. Moreover, since γ is a strictly decreasing function, $\{n_t\}_{t\in\mathcal{T}}$ is a strictly increasing sequence.

Suppose that $(m, t)\,\mathcal{R}^*(m^*, t^*)$. This implies that $u(m)\gamma(t) \geq u(m^*)\gamma(t^*)$. Therefore, $u(m) \geq u^*\gamma^*/\gamma(t) = u(n_t)$. By strict monotonicity of u, it must be that $m \geq n_t$. Analogously, we show that $(m^*, t^*)\,\mathcal{R}^*(m, t)$ implies $n_t \geq m$. □

Before we proceed with the proof Proposition 2.3, we need to introduce one additional notion. We will say that collection $\{(x^i, y^i)\}_{i\in I}$, where we denote $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, is a *dominant* sample of $\mathcal{R}^*$, if for all $m \in \mathcal{M}$ and $t \in \mathcal{T}$, we have

$$|\{i \in I : m^i \leq m\}| \geq |\{i \in I : n^i \leq m\}| \quad \text{and} \quad |\{i \in I : t^i \leq t\}| \leq |\{i \in I : s^i \leq t\}|.$$

We say that a sample is *strictly dominant*, whenever it is dominant, and there exists no subset $J \subseteq I$, such that for all $m \in \mathcal{M}$ and $t \in \mathcal{T}$, we have

$$|\{i \in J : m^i \leq m\}| \leq |\{i \in J : n^i \leq m\}| \quad \text{and} \quad |\{i \in J : t^i \leq t\}| \geq |\{i \in J : s^i \leq t\}|.$$

Clearly, whenever the sample is dominant and such subset exists, then $\{(x^i, y^i)\}_{i\in I\setminus J}$ is also a dominant sample. It is straightforward to show that for any dominant sample $\{(x^i, y^i)\}_{i\in I}$ such that either $|\{i \in I : m^i \leq m\}| > |\{i \in I : n^i \leq m\}|$, for some $m \in \mathcal{M}$, or $|\{i \in I : t^i \leq t\}| < |\{i \in I : s^i \leq t\}|$, for some $t \in \mathcal{T}$, has a strictly dominant sub-sample. We proceed with the proof of Proposition 2.3.

Proof of Proposition 2.3. We prove (i) $\Rightarrow$ (ii) by contradiction. Suppose that the set of observations is $\mathcal{O}$ is cyclically consistent, but there is a sample $\{(x^i, y^i)\}_{i\in I}$, where $x^i = (m^i, t^i)$ and $y^i = (n^i, s^i)$, such that for all $m \in \mathcal{M}$ and $t \in \mathcal{T}$, we have $|\{i \in I : m^i \leq m\}| \geq |\{i \in I : n^i \leq m\}|$ and $|\{i \in I : t^i \leq t\}| \leq |\{i \in I : s^i \leq t\}|$, where at least one inequality is strict. In particular, this implies that there exists a strictly dominant sub-sample $\{(x^j, y^j)\}_{j\in J}$, $J \subseteq I$.

Take any element (x^j, y^j) of the sub-sample. Since it is strictly dominant, it cannot be that $x^j \geq_X y^j$. Otherwise, we would have

$$|\{i \in \{j\} : m^i \leq m\}| \leq |\{i \in \{j\} : n^i \leq m\}| \quad \text{and} \quad |\{i \in \{j\} : t^i \leq t\}| \geq |\{i \in \{j\} : s^i \leq t\}|,$$

for all $m \in \mathcal{M}$ and $t \in \mathcal{T}$, which would violate that $\{(x^j, y^j)\}_{j\in J}$ is strictly dominant. On the other hand, since $\mathcal{O}$ is cyclically consistent, it cannot be that $y^j >_X x^j$. Otherwise, this would imply that $x^j \, \mathcal{R}^* y^j$ and $y^j >_X x^j$, which would violate the axiom. Hence, for any element (x^j, y^j) of the sample, x^j and y^j must be unordered with respect to the partial order $\geq_X$. Therefore, either (i) $m^j > n^j$ and $t^j > s^j$, or (ii) $m^j < n^j$ and $t^j < s^j$. Finally, since $\mathcal{E}$ is an anchored experiment, for any element of the sample, we have either $x^j = x^*$ and $y^j \neq x^*$, or $x^j \neq x^*$ and $y^j = x^*$. Take any pair (x^j, y^j) from the sample and consider the following claims.

Claim 1: If $x^j = x^*$, $n^j > (<)\ m^*$, and $s^j > (<)\ t^*$, then there is some (x^k, y^k) in the sample such that $m^k \geq (\leq)\ n^j$ and $t^k > (<)\ s^j$, or $m^k > (<)\ n^j$ and $t^k \geq (\leq)\ s^j$.

We prove the claim outside the brackets. Since the sample is dominant, it contains some (x^k, y^k) such that $t^k \geq s^j$. Since $\mathcal{E}$ is an anchored experiment, this implies that $y^k = x^*$. First, suppose that $t^k \geq s^j$ and $m^k < n^j$. Then $y^j >_X x^k$. However, since $x^k \, \mathcal{R}^* \, x^* \, \mathcal{R}^* y^j$, this would violate cyclical consistency. On the other hand, whenever $m^k = n^j$ and $t^k = s^j$, we have $x^k = y^j$ and $y^k = x^j$. In particular, this implies that for all $m \in \mathcal{M}$ and $t \in \mathcal{T}$, we have $|\{i \in \{j,k\} : m^i \leq m\}| = |\{i \in \{j,k\} : n^i \leq m\}|$ and $|\{i \in \{j,k\} : t^i \leq t\}| = |\{i \in \{j,k\} : s^i \leq t\}|$, which contradicts that $\{(x^j, y^j)\}_{j\in J}$ is strictly dominant. Therefore, it must be either $m^k \geq n^j$ and $t^k > s^j$, or $m^k > n^j$ and $t^k \geq s^j$. In order to prove the version in the brackets, note that the sample must contain some (x^k, y^k) such that $m^k \leq n^j$. The rest of the argument is analogous.

Claim 2: If $y^j = x^*$, $m^j > (<)\ m^*$, and $t^j > (<)\ t^*$, then there is some (x^k, y^k) in the sample such that $n^k \geq (\leq)\ m^j$ and $s^k > (<)\ t^j$, or $n^k > (<)\ m^j$ and $s^k \geq (\leq)\ t^j$.

We prove the claim analogously to Claim 1. Given Claims 1 and 2, there exists a sequence $\{x^k\}_{k=1}^K$ in $\mathcal{A}$ such that for any two subsequent elements $x^k = (m^k, t^k)$ and $x^{k+1} = (m^{k+1}, t^{k+1})$, we have $m^k \geq m^{k+1}$ and $t^k \geq t^{k+1}$, where at least one of the above inequalities is strict. Clearly, since $\mathcal{A}$ is finite, there exists the final element x^K of the sequence, for which there is no $y = (n, s)$ in $\mathcal{A}$ such that $m^K \geq n$ and $t^K \geq s$ (with at least one inequality being strict). However, by Claims 1 and 2, this contradicts the existence of a properly embedded sample of $\mathcal{R}^*$. Therefore, given that $\mathcal{O}$ obeys cyclical consistency, there is no dominant sample $\{(x^i, y^i)\}_{i\in I}$ of $\mathcal{R}^*$ such that, for some $m \in \mathcal{M}$ or $t \in \mathcal{T}$, we have $|\{i \in J : m^i \leq m\}| > |\{i \in J : n^i \leq m\}|$ or $|\{i \in J : t^i \leq t\}| < |\{i \in J : s^i \leq t\}|$. Hence, the dominance axiom is satisfied.

By Theorem 2.1, this implies that the set of observations $\mathcal{O}$ is rationalisable by a discounted utility function.

To show (ii) $\Rightarrow$ (iii), assume that (ii) holds. By Lemma B.6, there exists a strictly increasing sequence of numbers $\{n_t\}_{t\in\mathcal{T}}$ such that for any $x \in \mathcal{A}$, where $x = (m,t)$, whenever $x\,\mathcal{R}^* x^*$ then $m \geq n_t$, and $x^*\mathcal{R}^* x$ implies $m \leq n_t$. Take any number $u^* > 0$ and a discounting function $\gamma : \mathbb{N} \to (0,1]$, with $\gamma(0) = 1$. Let $\gamma^* := \gamma(t^*)$ and $u_t := u^*\gamma^*/\gamma(t)$. As γ is strictly decreasing, sequence $\{u_t\}_{t\in\mathcal{T}}$ is strictly increasing and positive.

Whenever $n_{\underline{t}} \neq 0$, set $n_0 = 0$, and let u_0 be any strictly positive number such that $u_0 < u_{\underline{t}}$. Otherwise, let $n_0 = n_{\underline{t}}$ and $u_0 = u_{\underline{t}}$. Denote $\mathcal{T}_0 := \mathcal{T} \cup \{0\}$. Finally, for all $t \in \mathcal{T}_0$ different from $\bar{t}$, let the immediate successor of t be $t^+ := \min\{s \in \mathcal{T} : s > t\}$. We define function $u : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$u(m) := \sum_{t\in\mathcal{T}_0} [u_t + \lambda_t(m - n_t)]\chi_{N_t}(m),$$

where $\lambda_j = (u_{t^+} - u_t)/(n_{t^+} - n_t)$ for all $t \neq \bar{t}$, $\lambda_{\bar{t}}$ is any strictly positive number, and $N_t := [n_t, n_{t^+})$, for all $t \neq \bar{t}$, while $N_{\bar{t}} := [n_{\bar{t}}, \infty)$. Clearly, the function is continuous and strictly increasing. Moreover, by construction, for all $t \in \mathcal{T}$, we have $u(n_t) = u_t$.

To complete this part of the proof, suppose that for some $x \in \mathcal{A}$, we have $x\,\mathcal{R}^* x^*$, where $x = (m,t)$. By assumption, it must be that $m \geq n_t$. By monotonicity of u, we have $u(m) \geq u(n_t) = u^*\gamma^*/\gamma(t)$, which implies that $u(m)\gamma(t) \geq u^*\gamma^* = u(m^*)\gamma(t^*)$. Analogously, we show that if $x^*\,\mathcal{R}^* x$ then $u(m^*)\gamma(t^*) \geq u(m)\gamma(t)$.

In order to complete the proof, note that implication (iii) $\Rightarrow$ (i) holds trivially, since any set rationalisable by a discounted utility function is rationalisable. $\square$

B.4 Proof of Theorem 2.2

In the following two lemmas we prove the sufficiency of the cumulative dominance axiom for rationalisation by a weakly present-biased discounted utility function.

Lemma B.7. *Let $\mathcal{O}$ obey the cumulative dominance axiom. There is a strictly increasing sequence $\{\phi_m\}_{m\in\mathcal{M}}$, a strictly decreasing sequence $\{\varphi_t\}_{t\in\mathcal{T}}$, and a strictly negative sequence $\{v_t\}_{t\in\mathcal{T}}$ of real numbers such that whenever $(m,t)\,\mathcal{R}^*(n,s)$ then $\phi_m + \varphi_t \geq \phi_n + \varphi_s$, while for all $t \in \mathcal{T}$, we have $\varphi_t + v_t(s-t) \leq \varphi_s$, for all $s \in \mathcal{T}$.*

Proof. Enumerate the elements of $\mathcal{R}^*$ such that it is equal to $\{(x^j, y^j)\}_{j\in J}$, where $x^j = (m^j, t^j)$ and $y^j = (n^j, s^j)$. In addition, define the immediate successor of t in $\mathcal{T}$ by $t^+ := \min\{s \in \mathcal{T} : s > t\}$, for all $t \in \mathcal{T}$ different from $\bar{t}$, while $\bar{t}^+ := \bar{t} + 1$.

For any $m \in \mathcal{M}$, let $\mu_m \in \{0,1\}^{|\mathcal{M}|}$ be a vector that takes the value of 1 at the coordinate corresponding to m, and 0 elsewhere. Moreover, for any $t \in \mathcal{T}$, let $\tau_t \in \{0,1\}^{|\mathcal{T}|}$ be a vector taking the value of $(t^+ - t)$ at the coordinate corresponding to t, and zero in all the remaining entries.

By $\mathbb{I}$ we denote a $|\mathcal{M}| + |\mathcal{T}|$ by $|\mathcal{M}| + |\mathcal{T}|$ identity matrix. Moreover, let B_1 be a $|J|$ times $|\mathcal{M}|$ matrix such that, for any $j \in J$, the j'th row of the matrix is equal to $\big(\sum_{k\le m^j} \mu_k - \sum_{k\le n^j} \mu_k\big)$. Similarly, let B_2 be a $|J|$ by $|\mathcal{T}|$ matrix where the j'th row is equal to $\big(\sum_{k\ge t^j} \tau_k - \sum_{k\ge s^j} \tau_k\big)$. Define B_3 as a $|\mathcal{T}|-1$ times $|\mathcal{M}|$ matrix with every entry equal to 0. Finally, B_4 is a matrix of dimensions $|\mathcal{T}| - 1$ by $|\mathcal{T}|$, where each column corresponds to one element in $\mathcal{T}$ and each row corresponds to an element in $\mathcal{T}\backslash\{\underline{t}\}$. Moreover, for each row corresponding to time-delay t, the entry in the column corresponding to t is equal to 1, while the entry in the column corresponding to t^+ is equal to -1. We set all the remaining entries to be equal to 0. We use the above matrices to construct a $|J| + |\mathcal{T}| - 1$ times $|\mathcal{M}| + |\mathcal{T}|$ matrix B, defined by

$$B := \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}.$$

We claim that whenever $\mathcal{O}$ obeys the cumulative dominance axiom, there exist vectors $\xi \in \mathbb{R}^{|\mathcal{M}|}$ and $\vartheta \in \mathbb{R}^{|\mathcal{T}|}$ such that $\mathbb{I} \cdot (\xi, \vartheta) \gg 0$ and $B \cdot (\xi, \vartheta) \ge 0$.

We prove the claim by contradiction. Suppose that $\mathcal{O}$ obeys the cumulative dominance axiom, but there are no such vectors. Theorem A.1 implies that there are some $\theta \in \mathbb{Z}^{|\mathcal{M}|+|\mathcal{T}|}$ and $\lambda \in \mathbb{Z}^{|J|+|\mathcal{T}|-1}$ such that $\theta \cdot \mathbb{I} + \lambda \cdot B = 0$, where $\theta > 0$ and $\lambda \ge 0$. Take any such vectors and denote $\lambda = (\overline{\lambda}, \underline{\lambda})$, where $\overline{\lambda} = (\lambda_j)_{j\in J}$ and $\underline{\lambda} = (\lambda_t)_{t\in\mathcal{T}\backslash\{\underline{t}\}}$, so that every coordinate of $\overline{\lambda}$ corresponds to a single element in $\mathcal{R}^*$, while each entry of $\underline{\lambda}$ corresponds to a time-delay $t \ne \underline{t}$. Moreover, let $\overline{B} := [B_1\ B_2]$, while $\underline{B} := [B_3\ B_4]$. Clearly, we have $\theta \cdot \mathbb{I} + \overline{\lambda} \cdot \overline{B} + \underline{\lambda} \cdot \underline{B} = 0$ $(\star)$.

Construct a sample $\{(x^i, y^i)\}_{i\in I}$ of $\mathcal{R}^*$ by taking λ_j copies of pair (x^j, y^j) from the directly revealed preference relation, for each $j \in J$. Since vector θ is strictly greater than zero, for condition $(\star)$ to hold, it must be that $\underline{\lambda} \cdot \underline{B} < -\overline{\lambda} \cdot \overline{B}$. In particular, this implies that $\sum_{i\in I} \sum_{k\le m^i} \mu_k \le \sum_{i\in I} \sum_{k\le n^i} \mu_k$. Hence, for all $m \in \mathcal{M}$, we have $|\{i \in I : m^i \le m\}| \ge |\{i \in I : m^i \le m\}|$.

Moreover, we require that $\underline{\lambda} \cdot \underline{B} \leq \sum_{i\in I}\sum_{k\geq s^i}\tau_k - \sum_{i\in I}\sum_{k\geq t^i}\tau_t$. Note that by definition, $\sum_{i\in I}\sum_{k\geq t^i}\tau_t = \int_t^{t^+}|\{i\in I : t^i \leq z\}|dz$, for all $t\in\mathcal{T}$. Hence, in particular, the initial condition implies that

$$\lambda_{\underline{t}} \leq \int_{\underline{t}}^{\underline{t}^+}|\{i\in I : s^i \leq z\}|dz - \int_{\underline{t}}^{\underline{t}^+}|\{i\in I : t^i \leq z\}|dz.$$

Since $\lambda_{\underline{t}} \geq 0$, we have $\int_{\underline{t}}^{\underline{t}^+}|\{i\in I : t^i \leq z\}|dz \leq \int_{\underline{t}}^{\underline{t}^+}|\{i\in I : s^i \leq z\}|dz$. Denote the immediate predecessor of t in $\mathcal{T}$ by $t^- := \max\{s\in\mathcal{T} : s < t\}$. Take any $t\in\mathcal{T}$, different from $\underline{t}$, and suppose that

$$\lambda_{t^-} \leq \int_{\underline{t}}^{t}|\{i\in I : s^i \leq z\}|dz - \int_{\underline{t}}^{t}|\{i\in I : t^i \leq z\}|dz. \tag{B.2}$$

For inequality $(\star)$ to hold, we require that

$$\lambda_t - \lambda_{t^-} \leq \int_{t}^{t^+}|\{i\in I : s^i \leq z\}|dz - \int_{t}^{t^+}|\{i\in I : t^i \leq z\}|dz.$$

Given the initial condition (B.2), this implies that

$$\lambda_t \leq \int_{\underline{t}}^{t^+}|\{i\in I : s^i \leq z\}|dz - \int_{\underline{t}}^{t^+}|\{i\in I : t^i \leq z\}|dz.$$

Since $\lambda_{t^+} \geq 0$, we have $\int_{\underline{t}}^{t^+}|\{i\in I : t^i \leq z\}|dz \leq \int_{\underline{t}}^{t^+}|\{i\in I : s^i \leq z\}|dz$. By induction, we conclude that the condition holds for all $t\in\mathcal{T}$.

Recall that $\theta > 0$. Therefore, for $(\star)$ to hold, it must be that at least one of the above inequalities is strict. However, this violates the cumulative dominance axiom. Therefore, there exist vectors ξ and ϑ such that $\mathbb{I}\cdot(\xi,\vartheta) \gg 0$ and $B\cdot(\xi,\vartheta) \geq 0$.

Take any pair (ξ,ϑ) satisfying the above system of inequalities. Define sequences $\{\xi_m\}_{m\in\mathcal{M}}$ and $\{\vartheta_t\}_{t\in\mathcal{T}}$ by $\xi_m := \mu_m\cdot\xi$ and $\vartheta_t := \tau_t\cdot\vartheta$ respectively. Clearly, both sequences are strictly positive. Moreover, whenever we have $(m,t)\,\mathcal{R}^*(n,s)$, then $\sum_{k\leq m}\xi_k + \sum_{k\geq t}\vartheta_k \geq \sum_{k\leq n}\xi_k + \sum_{k\geq s}\vartheta_k$. Define $\{\phi_m\}_{m\in\mathcal{M}}$ by $\phi_m := \sum_{k\leq m}\xi_k$. Given that $\{\xi_m\}_{m\in\mathcal{M}}$ is strictly positive, the above sequence is strictly increasing. Similarly, construct a strictly decreasing sequence $\{\varphi_t\}_{t\in\mathcal{T}}$ by $\varphi_t := \sum_{k\geq t}\vartheta_k$. By the previous argument, whenever $(m,t)\,\mathcal{R}^*(n,s)$ then $\phi_m + \varphi_t \geq \phi_n + \varphi_s$.

In order to complete the proof, define $v_t := -\vartheta_t/(t^+ - t)$, for all $t\in\mathcal{T}$. By definition of $\{\vartheta_t\}_{t\in\mathcal{T}}$, sequence $\{v_t\}_{t\in\mathcal{T}}$ is strictly negative and (weakly) increasing.

We need to show that for all $t \in \mathcal{T}$, we have $\varphi_t + v_t(s-t) \leq \varphi_s$, for all $s \in \mathcal{T}$. Clearly, the inequality holds whenever $s = t$. Assume that $s > t$. By definition of φ_t, we have

$$\begin{aligned}
\varphi_t + v_t(s-t) &= -\sum_{r\geq t}(r^+ - r)v_r + v_t(s-t) \\
&= -\sum_{r\geq t}(r^+ - r)v_r + v_t \sum_{s>r\geq t}(r^+ - r) \\
&= \sum_{s>r\geq t}(r^+ - r)(v_t - v_r) - \sum_{r\geq s}(r^+ - r)v_r \\
&\leq -\sum_{r\geq s}(r^+ - r)v_r \\
&= \sum_{r\geq s}\vartheta_r \\
&= \varphi_s,
\end{aligned}$$

where the inequality follows form the fact that $\{v_t\}_{t\in\mathcal{T}}$ is an increasing sequence. Using a similar argument, we can show that the condition holds for any $s < t$. □

Given the above lemma, we can show how to construct a weakly present-biased discounted utility function that rationalises the set of observations.

Lemma B.8. *Set $\mathcal{O}$ is rationalisable by a weakly present-biased discounted utility function, whenever there is a strictly increasing sequence $\{\phi_m\}_{m\in\mathcal{M}}$, a strictly decreasing sequence $\{\varphi_t\}_{t\in\mathcal{T}}$, and a strictly negative sequence $\{v_t\}_{t\in\mathcal{T}}$ of real numbers such that $(m,t)\,\mathcal{R}^*(n,s)$ implies $\phi_m + \varphi_t \geq \phi_n + \varphi_s$, while for all $t \in \mathcal{T}$, we have $\varphi_t + v_t(s-t) \leq \varphi_s$, for all $s \in \mathcal{T}$.*

Proof. Define function $\phi : \mathbb{R}_+ \to \mathbb{R}$ as in (B.1). Whenever $\underline{t} = 0$, construct a sequence $\{\tilde{\varphi}_t\}_{t\in\mathcal{T}}$, where $\tilde{\varphi}_t := \varphi_t - \varphi_{\underline{t}}$. Otherwise, let $\{\tilde{\varphi}_t\}_{t\in\mathcal{T}}$ be equal to $\{\varphi_t\}_{t\in\mathcal{T}}$. Moreover, denote $\tilde{\varphi}_0 = 0$ and $\mathcal{T}_0 := \mathcal{T} \cup \{0\}$. Define function $\bar{\varphi} : \mathbb{R} \to \mathbb{R}_-$, by

$$\bar{\varphi}(s) := \max_{t\in\mathcal{T}_0}\{\tilde{\varphi}_t + v_t(s-t)\}.$$

Note that, by definition of $\{\tilde{\varphi}_t\}_{t\in\mathcal{T}}$ and $\{v_t\}_{t\in\mathcal{T}}$, the above function is strictly decreasing, and convex. Hence, function $\vartheta(t) := \varphi(t) - \varphi(t+1)$ is also decreasing. Moreover, we have $\bar{\varphi}(t) = \tilde{\varphi}_t$, for all $t \in \mathcal{T}$. Finally, the above properties hold once we restrict the domain of $\bar{\varphi}$ to $\mathbb{N}$.

Define functions $u : \mathbb{R}_+ \to \mathbb{R}_+$ and $\gamma : \mathbb{N} \to (0,1]$ by $u := \exp(\phi)$ and $\gamma := \exp(\bar{\varphi})$ respectively. Clearly, the two functions are strictly monotone, while γ is log-convex with $\gamma(0) = 1$. Moreover, we have $v(m,t) := u(m)\gamma(t) = \exp(\phi_m + \tilde{\varphi}_t)$, for any $(m,t) \in \mathcal{A}$. Therefore, whenever $(m,t)\,\mathcal{R}^*(n,s)$ then $\phi_m + \tilde{\varphi}_t \geq \phi_n + \tilde{\varphi}_s$, which implies that $v(m,t) \geq v(n,s)$. □

Lemmas B.7 and B.8 imply that whenever the set of observation satisfies the cumulative dominance axiom, it is always possible to construct a utility function $u : \mathbb{R}_+ \to \mathbb{R}_+$ and a weakly present-biased discounting function $\gamma : \mathbb{N} \to (0,1]$ such that $v(m,t) := u(m)\gamma(t)$ rationalises $\mathcal{O}$. This concludes the proof of Theorem 2.2.

B.5 Proofs of Propositions 2.5 and 2.6

In the following section we present two lemmas that support the sufficiency of the strong cumulative dominance axiom for rationalisability by a quasi-hyperbolic discounted utility function. At the same time, the two results complete the proof of Proposition 2.5. In the second part of the section we prove Proposition 2.6.

Lemma B.9. *Whenever $\mathcal{O}$ obeys the strong cumulative dominance axiom for some $t' \in \mathcal{T}$, there exists a strictly increasing sequence $\{\phi_m\}_{m\in\mathcal{M}}$ and numbers $\hat{\beta}$, $\hat{\delta} < 0$ such that $(m,t)\,\mathcal{R}^*\,(n,s)$ implies $\phi_m + \min\{t,t'\}\hat{\beta} + t\hat{\delta} \geq \phi_n + \min\{s,t'\}\hat{\beta} + s\hat{\delta}$.*

Proof. Take any $t \in \mathcal{T}$ for which $\mathcal{O}$ obeys the strong cumulative dominance axiom. Enumerate the elements of $\mathcal{R}^*$ so that it is equal to $\{(x^j, y^j)\}_{j\in J}$, where $x^j = (m^j, t^j)$ and $y^j = (n^j, s^j)$. For any $m \in \mathcal{M}$, let $\mu_m \in \{0,1\}^{|\mathcal{M}|}$ be a vector that takes the value of 1 at the coordinate corresponding to m, and 0 everywhere else. Let $\mathbb{I}$ denote a $|\mathcal{M}| + 2$ by $|\mathcal{M}| + 2$ a diagonal matrix, where for the first $|\mathcal{M}|$ rows the corresponding entries are equal to 1, while for the last two are equal to -1. Let B be a $|J|$ times $|\mathcal{M}| + 2$ matrix such that, for any $j \in J$, its j'th row is equal to $(\sum_{k\leq m^j}\mu_k - \sum_{k\leq n^j}\mu_k,\ t^j - s^j,\ \min\{t^j,t'\} - \min\{s^j,t'\})$. We claim that there exists a vector $\xi \in \mathbb{R}$ and real numbers $\hat{\beta}$ and $\hat{\delta}$ such that $\mathbb{I}\cdot(\xi,\hat{\delta},\hat{\beta}) \gg 0$ and $B\cdot(\xi,\hat{\delta},\hat{\beta}) \geq 0$.

We prove the claim by contradiction. Suppose that there are no such vectors. By Theorem A.1 (see Appendix A), there is some $\theta \in \mathbb{Z}^{|\mathcal{M}|+2}$ and $\lambda \in \mathbb{Z}^{|J|}$ such that $\theta\cdot\mathbb{I} + \lambda\cdot B = 0$ $(\star)$, where $\theta > 0$ and $\lambda \geq 0$. Take any such vectors and denote $\lambda = (\lambda_j)_{j=1}^J$. Construct sample $\{(x^i, y^i)\}_{i\in I}$ by taking λ_j copies of pair (x^j, y^j) from of $\mathcal{R}^*$, for all $j \in J$. Since vector θ is strictly greater than zero, for condition $(\star)$ to hold, it must be that $\sum_{i\in I}\sum_{k\leq m^i}\mu_k \leq \sum_{i\in I}\sum_{k\leq n^i}\mu_m$, as well as $\sum_{i\in I} t^i \geq \sum_{i\in I} s^j$, and

$\sum_{i\in I}\min\{t^i,t'\} \geq \sum_{i\in I}\min\{s^i,t'\}$, with at least one inequality being strict. However, this contradicts the strong cumulative dominance axiom.

Take any ξ and $\hat{\delta}$, $\hat{\beta}$ satisfying $\mathbb{I}\cdot(\xi,\hat{\delta},\hat{\beta}) \gg 0$ and $B\cdot(\xi,\hat{\delta},\hat{\beta}) \geq 0$. Define sequence $\{\xi_m\}_{m\in\mathcal{M}}$ by $\xi_m := \mu_m\cdot\xi$. By construction of matrix B, whenever $(m,t)\,\mathcal{R}^*(n,s)$ then $\sum_{k\leq m}\xi_k + t\hat{\delta} + \min\{t,t'\}\hat{\beta} \geq \sum_{k\leq n}\xi_k + s\hat{\delta} + \min\{s,t'\}\hat{\beta}$. Define sequence $\{\phi_m\}_{m\in\mathcal{M}}$ by $\phi_m := \sum_{k\leq m}\xi_k$. Since $\{\xi_m\}_{m\in\mathcal{M}}$ is strictly positive, $\{\phi_m\}_{m\in\mathcal{M}}$ is strictly increasing and satisfies the property specified in the lemma. □

Given the above lemma, the following result completes the sufficiency part of Proposition 2.5.

Lemma B.10. *Set $\mathcal{O}$ is rationalisable by a quasi-hyperbolic discounting function if there is some $t' \in \mathcal{T}$, a strictly increasing sequence $\{\phi_m\}_{m\in\mathcal{M}}$ and numbers $\hat{\beta},\hat{\delta} < 0$ such that $(m,t)\,\mathcal{R}^*\,(n,s)$ implies $\phi_m + t\hat{\delta} + \min\{t,t'\}\hat{\beta} \geq \phi_n + s\hat{\delta} + \min\{s,t'\}\hat{\beta}$.*

Proof. Construct function $\phi : \mathbb{R}_+ \to \mathbb{R}$ as in (B.1). Recall that the function is continuous and strictly increasing. Moreover, for any $m \in \mathcal{M}$, we have $\phi(m) = \phi_m$. Define function $u : \mathbb{R}_+ \to \mathbb{R}_+$ by $u := \exp(\phi)$. Moreover, let $\delta := \exp(\hat{\delta})$, $\beta := \exp(\hat{\beta})$, and $t^\circ := t'$. Define function $\gamma : \mathbb{N} \to (0,1]$ by $\gamma(t) := \beta^t\delta^t$, whenever $t < t^\circ$, and $\gamma(t) := \beta^{t^\circ}\delta^t$ otherwise. Clearly, we have $\gamma(0) = 1$. Moreover, for any prize-time pair $(m,t) \in \mathcal{A}$, we have $v(m,t) := u(m)\gamma(t) = \exp(\phi_m + t\hat{\delta} + \min\{t,t^\circ\}\hat{\beta})$. Therefore, whenever $(m,t)\,\mathcal{R}^*(n,s)$ holds, then $v(m,t) \geq v(n,s)$. The proof is complete. □

We complete this section with the argument supporting Proposition 2.6 which is presented in Section 2.4.3.

Proof of Proposition 2.6. We can show the necessity of the conditions stated in the proposition analogously to the quasi-hyperbolic case. In order to show sufficiency, enumerate the elements of $\mathcal{R}^*$ such that it is equal to $\{(x^j,y^j)\}_{j\in J}$, where $x^j = (m^j,t^j)$ and $y^j = (n^j,s^j)$. For each $m \in \mathcal{M}$, let $\mu_m \in \{0,1\}^{|\mathcal{M}|}$ be a vector with all entries equal to zero, apart from the one corresponding to m equal to 1. Moreover, denote the immediate successor of m in $\mathcal{M}$ by $m^+ := \min\{n \in \mathcal{M} : n > m\}$, for all $m \neq \overline{m}$. Let A be a $|\mathcal{M}|$ by $|\mathcal{M}|+1$ matrix constructed as follows. Each of the first $|\mathcal{M}|-1$ rows is equal to $(\mu_{m^+}-\mu_m,0)$, while the $|\mathcal{M}|+1$'th entry of the $|\mathcal{M}|$'th row is equal to -1. Let B be a $|J|$ by $|\mathcal{M}|+1$ matrix, where for each $j \in J$, its j'th row is equal to $(\mu_{m^j} - \mu_{n^j}, t^j - s^j)$. We claim that whenever *mathcalO* obeys the condition stated in the proposition, there is a vector $\phi \in \mathbb{R}^{|\mathcal{M}|}$ and a number $\hat{\delta}$ such that $A \cdot (\phi, \hat{\delta}) \gg 0$ and $B \cdot (\phi, \hat{\delta}) \geq 0$.

We prove the claim by contradiction. Suppose that $\mathcal{O}$ obeys the condition, but the above system of inequalities has no solution. By Theorem A.1, there exists $\theta \in \mathbb{Z}^{|\mathcal{M}|}$ and $\lambda \in \mathbb{Z}^{|J|}$, with $\theta > 0$ and $\lambda \geq 0$, such that $\theta \cdot A + \lambda \cdot B = 0$ $(\star)$. Take any such vectors and denote $\lambda = (\lambda_j)_{j=1}^{J}$. Construct a sample $\{(x^i, y^i)\}_{i \in I}$ of $\mathcal{R}^*$, by taking λ_j times each pair (x^j, y^j). By definition of matrices A and B, whenever $(\star)$ is satisfied, the above sample can be partitioned into subsets $\{(x^k, y^k)\}_{k \in K}$ of $\mathcal{R}^*$ such that, for all $m \in \mathcal{M}$, we have $|\{k \in K : m^k \leq m\}| \geq |\{k \in K : n^k \leq m\}|$. Moreover, it must be that $\sum_{i \in I} t^i \geq \sum_{i \in I} s^i$. Finally, as $\theta > 0$, at least one of the above inequalities must be strict. However, this contradicts the condition stated in the axiom.

Take any ϕ and $\hat{\delta}$ such that $A \cdot (\phi, \hat{\delta}) \gg 0$ and $B \cdot (\phi, \hat{\delta}) \geq 0$, and define sequence $\{\phi_m\}_{m \in \mathcal{M}}$ by $\phi_m := \mu_m \cdot \phi$. Clearly, it is strictly increasing, while $(m,t)\,\mathcal{R}^*(n,s)$ implies $\phi_m + \hat{\delta}t \geq \phi_n + \hat{\delta}s$. Construct function $\phi : \mathbb{R}_+ \to \mathbb{R}$ as in (B.1), which is continuous and strictly increasing and, for any $m \in \mathcal{M}$, we have $\phi(m) = \phi_m$. Define function $u : \mathbb{R}_+ \to \mathbb{R}_+$ by $u := \exp(\phi)$, and let $\delta = \exp(\hat{\delta})$. Clearly, for any $(m,t) \in \mathcal{A}$, we have $v(m,t) := \delta^t u(m) = \exp(\phi_m + t\hat{\delta})$. Hence, $(m,t)\,\mathcal{R}^*(n,s)$ implies $v(m,t) \geq v(n,s)$. The proof is complete. □

Bibliography

Afriat, S. N. (1972): "Efficiency estimation of production functions," *International Economic Review*, 13(3), 568–598.

Aliprantis, C. D., and K. C. Border (2006): *Infinite dimensional analysis: A hitchhiker's guide*. Springer Berlin, 3rd edn.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2008): "Eliciting risk and time preferences," *Econometrica*, 76(3), 583–618.

Balbus, L., K. Reffett, and L. Woźny (2014): "Computing time-consistent Markov policies for quasi-hyperbolic consumers under uncertainty," *International Journal of Game Theory*, forthcoming.

Barro, R. J. (1999): "Ramsey meets Laibson in the neoclassical growth model," *Quarterly Journal of Economics*, 114(4), 1125–1152.

Benhabib, J., A. Bisin, and A. Schotter (2010): "Present-bias, quasi-hyperbolic discounting, and fixed costs," *Games and Economic Behavior*, 69(2), 205–223.

Border, K. C. (1992): "Revealed preference, stochastic dominance, and the expected utility hypothesis," *Journal of Economic Theory*, 56, 20–42.

Brown, D. J., and C. Calsamiglia (2007): "The nonparametric approach to applied welfare analysis," *Economic Theory*, 31(1), 183–188.

Chabris, C. F., D. Laibson, C. L. Morris, J. P. Schuldt, and D. Taubinsky (2008): "Individual laboratory-measured discount rates predict field behavior," *Journal of Risk and Uncertainty*, 37(2), 237–269.

——— (2009): "The allocation of time in decision-making," *Journal of the European Economic Association*, 7(2-3), 628–637.

Chambers, C., and F. Echenique (2009): "Profit maximization and supermodular technology," *Economic Theory*, 40(2), 173–183.

Coller, M., and M. Williams (1999): "Eliciting individual discount rates," *Experimental Economics*, 2(2), 107–127.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2012): "Interpreting time horizon effects in inter-temporal choice," Discussion Paper No. 6385, IZA.

Echenique, F., T. Imai, and K. Saito (2014): "Testable implications of quasi-hyperbolic and exponential discounting," Caltech SS Working Paper 1388.

Fishburn, P. C., and A. Rubinstein (1982): "Time preference," *International Economic Review*, 23(3), 677–94.

Fourier, J. (1826): "Solution d'une question perticulière du calcul des inégalités," *Nouveau Bulletin des Sciences par la Sociéeté philomatique de Paris*, p. 317–319.

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time discounting and time preference: A critical review," *Journal of Economic Literature*, 40(2), 351–401.

Harris, C., and D. Laibson (2001): "Dynamic choices of hyperbolic consumers," *Econometrica*, 69(4), 935–57.

Harrison, G. W., M. I. Lau, and M. B. Williams (2002): "Estimating individual discount rates in Denmark: A field experiment," *American Economic Review*, 92(5), 1606–1617.

Herings, P. J.-J., and K. Rohde (2006): "Time-inconsistent preferences in a general equilibrium model," *Economic Theory*, 29(3), 591–619.

——— (2008): "On the completeness of complete markets," *Economic Theory*, 37(2), 171–201.

KIRBY, K. N., AND N. N. MARAKOVIC (1964): "Delay-discounting probabilistic rewards: Rates decrease and amounts increase," *Psychonomic Bulletin & Review*, 3, 100–104.

KIRBY, K. N., N. M. PETRY, AND W. K. BICKEL (1999): "Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls," *Journal of Experimental Psychology: General*, 128, 78–87.

KRAFT, C. H., J. W. PRATT, AND A. SEIDENBERG (1959): "Intuitive probability on finite sets," *Annals of Mathematical Statistics*, 30, 408–419.

LAIBSON, D. (1997): "Golden eggs and hyperbolic discounting," *Quarterly Journal of Economics*, 112(2), 443–77.

LUTTMER, E. G. J., AND T. MARIOTTI (2003): "Subjective discounting in an exchange economy," *Journal of Political Economy*, 111(5), 959–989.

——— (2006): "Competitive equilibrium when preferences change over time," *Economic Theory*, 27(3), 679–690.

——— (2007): "Efficiency and equilibrium when preferences are time-inconsistent," *Journal of Economic Theory*, 132(1), 493–506.

MAS-COLELL, A., M. D. WHINSTON, AND J. R. GREEN (1995): *Microeconomic theory*. Oxford University Press.

MEYER, M., AND B. STRULOVICI (2013): "The supermodular stochastic ordering," Discussion Papers 1563, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

MILGROM, P., AND J. ROBERTS (1990): "The Economics of modern manufacturing: Technology, strategy, and organization," *American Economic Review*, 80(3), 511–28.

——— (1994): "Complementarities and systems: Understanding Japanese economic organization," *Estudios Económicos*, 9(1), 3–42.

——— (1995): "Complementarities and fit strategy, structure, and organizational change in manufacturing," *Journal of Accounting and Economics*, 19(2-3), 179–208.

Milgrom, P., and C. Shannon (1994): "Monotone comparative statics," *Econometrica*, 62(1), 157–180.

Motzkin, T. S. (1951): "Two consequences of the transposition theorem on linear inequalities," *Econometrica*, p. 184–185.

Negishi, T. (1960): "Welfare economics and existence of equilibrium for a competitive economy," *Metroeconomica*, 12, 92–97.

Nishimura, H., E. A. Ok, and J. K.-H. Quah (2013): "A unified approach to revealed preference theory: The case of rational choice," Economics Series Working Papers 686, University of Oxford, Department of Economics.

Noor, J. (2011): "Intertemporal choice and the magnitude effect," *Games and Economic Behavior*, 72(1), 255–270.

Ok, E. A., and Y. Masatlioglu (2007): "A theory of (relative) discounting," *Journal of Economic Theory*, 137(1), 214–245.

Pollak, R. A. (1968): "Consistent planning," *Review of Economic Studies*, 35, 201–208.

Rochet, J. C. (1987): "A necessary and sufficient condition for rationalizability in a quasilinear context," *Journal of Mathematical Economics*, 16, 191–200.

Rockafellar, T. R. (1970): *Convex analysis.* Princeton University Press.

Sákovics, J. (2013): "Revealed cardinal preference," *Economic Theory Bulletin*, 1, 39–45.

Scott, D. (1964): "Measurement structures and linear inequalities," *Journal of Mathematical Psychology*, 1, 233–247.

Stoer, J., and C. Witzgall (1970): *Convexity and optimization in finite dimensions I.* Springer Verlag, Berlin.

Strotz, R. H. (1955): "Myopia and inconsistency in dynamic utility maximization," *Review of Economic Studies*, 23(3), 165–180.

Topkis, D. M. (1978): "Minimizing a submodular function on a lattice," *Operations Research*, 26(2), 305–321.

——— (1995): "Comparative statics of the firm," *Journal of Economic Theory*, 67, 370–401.