ESSAYS IN FINANCIAL ECONOMICS
by
AYMAN M. HINDY

B.Sc., Civil Engineering, Cairo University, June 1983
M.Sc., Civil Engineering, MIT, February 1987

Submitted to the Sloan School of Management
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

at the

Massachusetts Institute of Technology

September 1990

©Ayman M. Hindy, 1990. All rights reserved

The author hereby grants to MIT permission to reproduce and to
distribute copies of this thesis document in whole or in part.

Signature of Author

Sloan School of Management
June 30, 1990

Certified by

Chi-fu Huang
Professor of Finance
Thesis Supervisor

Accepted by

James B. Orlin
Chairman, Department Committee
ESSAYS IN FINANCIAL ECONOMICS
by
AYMAN M. HINDY

Submitted to the Sloan School of Management On June 30, 1990
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Management.

Abstract

This dissertation is composed of four independent essays. In Essay I, we report an exploratory study of the process of price formation in a speculative market in the absence of liquidity traders. Traders exchange a futures contract because they interpret information differently. We formulate trading as a sequence of anonymous double auctions and embed the price formation mechanism in a large economy from which random samples of orders are drawn for execution on the trading floor. We introduce a notion of bounded rationality in which traders use approximate models of market response in forming their bids. We justify this approach on the basis of a complexity argument and defend it on the grounds of a robustness criterion. We prove existence of a perfect equilibrium in the sequential anonymous auctions game, and show that the equilibrium has a "no-regret" property. After learning the market price, a trader regrets neither the bid that he made nor the position that he holds. We show that trading volume is related to changes in the distribution of information in the economy. We also show that volume and expected change in price are related to two different attributes of the pattern of private information flow. Fundamentally, no particular relationship between the time series of these variables is always valid for all futures contracts. This point is emphasized by an example.

In Essay II, we study the problem of optimal consumption choice in continuous time under certainty for a class of utility functions that capture the notion that consumptions at nearby dates are almost perfect substitutes. The class we consider excludes all time-additive and almost all the non time-additive utility functions used in the literature. We provide necessary and sufficient conditions for a consumption policy to be optimal. Furthermore, we demonstrate our general theory by solving in a closed form the optimal consumption policy for a particular felicity function. The optimal policy in our solution consists of a (possible) initial "gulp" of consumption, or an initial period of no consumption, followed by consumption at the rate that maintains a constant ratio of wealth to average past consumption.
In Essay III, we study the problem of optimal consumption and portfolio choice in continuous time under uncertainty for a class of utility functions that capture the notion that consumptions at nearby dates are almost perfect substitutes. The class we consider excludes all time-additive and almost all the non time-additive utility functions used in the literature. We provide sufficient conditions for a consumption and portfolio policy to be optimal. Furthermore, we demonstrate our general theory by solving in a closed form the optimal consumption and portfolio policy for a particular felicity function when the prices of the assets follow a geometric Brownian motion process. The optimal consumption policy in our solution consists of a possible initial “gulp” of consumption followed by a process of accumulated consumption with singular sample paths. In almost all states of nature, the agent consumes at uncountably infinite number of moments. However, the set of all times at which consumption occurs has Lebesgue measure zero.

Finally, in Essay IV, we survey the literature on dynamic choice theory and discuss measures of risk aversion in a dynamic framework. We summarize the difficulties of measuring attitudes towards risk when utility is derived from many commodities. We present Kreps and Porteus (1978) development of preferences over temporal lotteries and Selden’s (1978) ordinal certainty equivalent approach. We survey recent applications to asset pricing, notably Epstein and Zin (1989) contributions, and we provide a critique of their interpretations. Finally, we point out directions for further research.

Thesis Committee:
Chi-fu Huang, Professor of Finance, Sloan School of Management, MIT, Chairman.
David Scharfstein, Associate Professor of Finance, Sloan School of Management, MIT.
John Heaton, Assistant Professor of Finance and Economics, Sloan School of Management, MIT.
Acknowledgement

Words alone cannot express my deep gratitude to Chi-fu Huang, my dissertation advisor. Nonetheless, I will try. Chi-fu is a unique combination of great intellects and genuine caring for those persons for whom he has some responsibility. I was very fortunate to be one of these people. To me, Chi-fu is not only a world class advisor, he is also a dear friend, an influential mentor, a source of endless inspiration and a stimulating coauthor. As if this was not enough for him, Chi-fu cannot help but be a source of great fun even during the most stressful times. Chi-fu, thank you very much.

I am also grateful to many people at MIT for helpful discussions. Both David Scharfstein and John Heaton were generous with their time and comments and agreed to serve on my thesis committee. Avraham Beja, Andy Lo and Jean-Luc Vila had a lot to say about my work in a very constructive and pleasant manner.

My experience at MIT was enriched through interaction with many colleagues. My officemates, Anne Beatty and Lisa Meulbroek, were a source of comfort and support. Neil Pearson never failed me when I sought his advice. My friend Hua He was always there to read, comment and discuss. He also managed to find an error in an earlier version of Essay II.

Two dear friends have always had a great impact on my career. Hisham El-Sherif is a source of continuous encouragement and motivation, and Mohsen Baligh never fails to impress me by his wisdom and insights.

Last but not least, all members of my family, especially Faten, made so many contributions to my life during my graduate school days that I feel that this dissertation is theirs more than it is mine. It is only natural that I dedicate this work to my family.
Introduction

This dissertation is composed of four essays that deal with three different subjects. In Essay I, I investigate the topic of price formation in large anonymous markets. Essays II and III, which are coauthored with Chi-fu Huang, report an analysis of the problem of optimal consumption and portfolio choice in continuous time for a class of agents with non time-additive utility functions. In Essay IV, I study the issue of risk aversion in dynamic choice circumstances.

The research reported in Essay I is based on the view that one important function of financial markets is collecting and transmitting information. The prices established in the course of trading reflect the private information of different participants. This function of the markets has traditionally been analyzed using concepts of rational expectations equilibria; see, for example, Grossman (1976) and Grossman and Stiglitz (1980). An important premise behind the work in Essay I is that perhaps there is room for better understanding of markets as means of aggregating information if we model explicitly the process of price formation. Based on this thesis, I analyze a sequence of double auctions which is a price formation mechanism actually used in the trading pits of futures exchanges and, to some extent, in screen trading in over-the-counter markets for stocks.

Modeling price formation explicitly is not new. Many analyses of market microstructure; see, for example, Kyle (1985) and Admati and Pfleiderer (1988), account for the detailed steps of price determination. The analysis in Essay I is different from most models in the existing market microstructure literature because no "noise" or "liquidity" traders participate in the double auction. I consider the model in Essay I as a view of market dynamics which complements the existing models of market microstructure. In addition, it provides a description of an environment which might be a better approximation to some actual markets.

The main conclusion of Essay I is that modeling price formation in a large market without noise traders can be done, albeit at some conceptual cost. The price I had to pay is accepting "approximate" rather than "exact" solutions to the trading game. On the positive side, I introduce a notion of "robustness" of approximate models which has a behavioral interpretation and which could prove useful for other applications. There are two main results in Essay I. First, I prove existence of a perfect equilibrium in a sequence of double auctions. This result is a contribution to the literature on double auctions which happen to be very difficult to analyze. Second, I show that the changes in prices and trading volume are fundamentally unrelated. Further characterization of the way the distribution of information among traders changes over time is required to make predictions about the relation between price change and volume.

Essays II and III fit into the literature dealing with optimal consumption and portfolio choice in continuous time which was pioneered by Merton (1971). In Merton's original formulation and in subsequent work; see, for example, Merton (1973), Breeden (1979,
1986), and Cox, Ingersoll and Ross (1985), an agent maximizes a time-additive utility of consumption rates. This specification of preferences proved to be tractable and produced useful insights.

The assumption of time-additive utility, however, is very strong and economically not intuitive since it implies that consumption at one time has absolutely no effect on the instantaneous satisfaction at nearby times. Many researchers have questioned this assumption. In addition, equilibrium models based on this specification did not perform well empirically. The failure of the time-additive specification has manifested itself in many contexts, for example in the equity premium puzzle pointed out by Mehra and Prescott (1985).

The response in the literature to the deficiencies of the time-additive model has been along two lines. Some researchers; see, for example, Constantinides (1989), Sundaresan (1989) and Heaton (1990), have proposed some alternative specifications of utility which seem to capture the effects of consumption at one time on satisfaction at other nearby times. Another approach, taken by Huang and Kreps (1989) and Hindy and Huang (1989), starts from more fundamental axioms about "reasonable" intertemporal preferences for consumption in continuous time. This approach leads to a formal structure within which many specific functional forms of utility can be analyzed. Huang and Kreps (1989) and Hindy and Huang (1989) find out that many functional forms in the literature of non time-additive utility do not capture some intuitive notions about substitutability of consumption between two nearby times.

Essays II and III build on the work of Huang and Kreps (1989) and Hindy and Huang (1989). We study the problem of optimal consumption and portfolio choice, under certainty in Essay II and under uncertainty in Essay III, for a class of utility functions proposed by Hindy and Huang (1989). The class we consider excludes all the time-additive and almost all the non time-additive utilities in the literature. We provide sufficient conditions for optimality and show that consumption in such a model occurs only periodically and not continually as most models suggest. In addition, Essay III contains some novel analysis of a nonstandard optimal control problem. In constructing a solution to the control problem, we solve a partial differential equation using probabilistic techniques. In addition, the analysis in Essay III exploits some beautiful connections between analytic and probabilistic concepts.

Essay IV grew naturally from Essay III. The agent in Essay III has an instantaneous power utility function of average past consumption. In solving for the optimal investment rule, we find that he follows a constant proportion policy. The proportions of wealth invested in the risky assets do not change with time. However, the proportions invested in the risky assets are more than those for an agent with the same power instantaneous utility function which is rather defined on current consumption rate and who faces the same investment opportunity. In addition, if one computes the standard Arrow-Pratt measure of risk aversion using the indirect utility function, one finds that the agent with the non time-additive preferences is less risk averse. These observations lead naturally to
the question of how to measure risk aversion in dynamic choice settings and how changes in intertemporal preferences produce changes in attitudes towards risks.

In Essay IV, I first survey the literature on choice among lotteries with prizes in a multi-dimensional commodity space. The central message there is that there is an intimate connection between preferences on sure prizes and preferences on gambles on wealth. Assumptions made on one set of preferences impose restrictions on the other type of preferences. I then specialize this discussion to models of dynamic choice when the prizes are streams of future consumption. I present different ways of modeling this situation and provide a critique of some of them. The discussion then points out naturally to directions for further research.

Although, the topics of the four essays of the dissertation are different, their material is quite close. They all deal, in some form or another, with the workings of financial markets. Financial markets are complex institutions that have many facets and perform many functions. The information gathering, aggregation and transmittal function of these markets is dealt with in Essay I. Essays II, III and IV are related to financial markets in their function as a vehicle for risk sharing and intertemporal redistribution of consumption. The focus on financial markets from different angles is the unifying theme behind the topics studied in this dissertation.

References


Essay I

AN EQUILIBRIUM MODEL OF FUTURES MARKETS DYNAMICS

Abstract

In this essay, we report an exploratory study of the process of price formation in a speculative market in the absence of liquidity traders. Traders exchange a futures contract because they interpret information differently. We formulate trading as a sequence of anonymous double auctions and embed the price formation mechanism in a large economy from which random samples of orders are drawn for execution on the trading floor. We introduce a notion of bounded rationality in which traders use approximate models of market response in forming their bids. We justify this approach on the basis of a complexity argument and defend it on the grounds of a robustness criterion. We prove existence of a perfect equilibrium in the sequential anonymous auctions game, and show that the equilibrium has a “no-regret” property. After learning the market price, a trader regrets neither the bid that he made nor the position that he holds. We show that trading volume is related to changes in the distribution of information in the economy. We also show that volume and expected change in price are related to two different attributes of the pattern of private information flow. Fundamentally, no particular relationship between the time series of these variables is always valid for all futures contracts. This point is emphasized by an example.
1 Introduction and Summary

Perhaps there is no better symbol of competitive price formation mechanisms than the open outcry auctions conducted continuously in the trading pits of futures exchanges. Competitive bidding by floor brokers executing orders on behalf of customers outside the exchange determines prices. Soon, these customers will be able to place their bids to the market directly via automated trading systems. Typically, these customers think of the market as an entity with more or less known patterns of response. They always ask: Where is the market going? and How will the market respond to this piece of news? They also plan their trading according to their expectations about market behavior.

Recent models of market microstructure depict a large portion of these customers as "noise" or "liquidity" traders with a passive role in price determination. For example, in Kyle (1985) the demands of liquidity traders are independent of the expected market price. In Admati and Pfleiderer (1988) and (1989), liquidity traders are again passive in the sense that their reservation values are independent of the market maker's quoted bid and ask prices. Moreover, liquidity traders are expected to realize net losses to compensate the market maker for his expected loss to informed traders. In Glosten and Milgrom (1985), liquidity traders revise their expectations of the value of the asset after observing the dealer's quotes. They can only submit market orders, however, and whether an equilibrium actually exists is unknown.

These models provide us with valuable insights about some aspects of market microdynamics. They explain price response to the order arrival process and rationalize some observed intraday patterns of trading in markets with designated market makers. The image of "liquidity" traders, however, as passive agents who do not attempt to learn from the prices at which they transact and who consistently lose money is rather disquieting. This is particularly so when one considers a speculative market for relatively short-lived assets with high volatility. The assumed liquidity trader is a person or an institution who holds the assets for intertemporal allocation of resources and who for either predictable or unpredictable reasons decides to liquidate his position. Such a picture is convincing when the asset in question can function as a store of value for such a trader. Futures contracts hardly fit this description.
On the other hand, we can construct models in which all traders make inferences from prices using the rational expectations framework. Prices in this framework play the dual role of determining the budget constraint of traders as well as providing information about private signals. There are a number of problems with rational expectations equilibria as discussed by Milgrom (1981b), Bray and Kreps (1987), Dubey, Geanakopolos and Shubik (1987), among others. More important, rational expectations models do not explicitly analyze the process by which prices are formed.

This essay attempts to provide a complementary view to the existing microstructure literature. We report here an exploratory study in which we try to analyze an explicit price formation mechanism in a large market in the absence of “noise” traders. Specifically, we ask three questions. First: Could we formulate and analyze a model of price formation in a large market without introducing noise traders? Second: How could we capture the notion that in anonymous trading, each trader acts strategically against the “market”? What are the assumptions required for a tractable and potentially useful analysis of this notion and are these assumptions justifiable? Third: How would the patterns of price change and volume we get from analyzing such a model relate to those predicted by the existing literature?

Addressing the first question, we construct an equilibrium model of dynamic trading in an environment of heterogeneous information using an explicit price formation mechanism. The market we study is a futures market, an example of markets mainly driven by information flow, and the price formation mechanism we focus on is double auctions, a form widely used in actual futures markets.

The formulation we use accommodates a micro view of the trading floor as the place where price is determined as well as a global view in which the demands and information of all potential traders are explicitly recognized. We explicitly distinguish between the Economy and the Market. The Economy is the total population of potential traders, whereas the Market is a group of traders randomly selected from the Economy. Borrowing from the microstructure literature, we recognize the fact that markets have limited trading capacity and that trading takes time. At any point in time, the orders on the trading floor are just a representative sample of reservation prices of agents in the Economy. The reservation prices, however, are equilibrium reservation prices since they emerge
from the optimizing behavior of each agent using both private and public information. Thus, the market clearing price at any time is an equilibrium price in the sense that it comes from optimum bidding behavior, and it is also a market determined price in the sense that it depends on the actual design of the market. We can conceptually separate the determinants of prices into two major components: economy-wide related factors such as dissemination of private information and market-specific factors such as market capacity and order handling procedures. This distinction should prove useful for later work.

To address the second question about modeling the strategic behavior of traders, we need a game-theoretic approach. Recent contributions to the problem of price formation have relied on applying game theory to finely detailed models of trading processes. See, for example, Wilson (1987a). Such applications have been limited, however, to studying strategic interactions in markets with few traders. When one tries to apply the techniques of game theory to complex markets with a large number of participants, one encounters the following conceptual difficulty. At the foundations of game-theoretic models is the assumption that the structure of the game is common knowledge. One assumes that the number and identity of the players as well as the strategic options open to each are common knowledge. Moreover, in games of incomplete information, one assumes that the distributions of the preferences of players and their private information are also common knowledge. This common knowledge assumption seems too strict, since in large anonymous markets the identities of market participants against whom one trader strategically moves is usually not known, not to mention their preferences and information.

Wilson (1987a) writes

*I foresee the progress of game theory as depending on successive reductions in the base of common knowledge required to conduct useful analyses of practical problems. Only by repeated weakening of common knowledge assumptions will the theory approximate reality.*

There is, to be sure, a strategic aspect to the behavior of market participants. This strategic behavior, however, is not directed towards a particular known group of traders, but rather towards a large nonpersonal entity that one refers to as The Market. The
concept of *anonymous games*, Jovanovic and Rosenthal (1988), captures the features of strategic behavior without requiring the assumption of common knowledge of detailed characteristics of each player. An anonymous game is a non-cooperative game with a continuum of players possessing the feature that a player’s payoff depends on his opponents’ actions only through their distribution. In an anonymous game, if half the opponents take one action and half take another, the identities of those in each pool makes no difference to the player. The base of common knowledge is thus reduced to knowledge about the distributions of possible player characteristics, in substitute for a detailed description of the characteristics of each player. In this essay we use anonymous sequential games as the framework of the trading model.

An individual participating in this market needs to form a model about the aggregate behavior of individuals for inference and prediction purposes. Most models of dynamic trading in a homogeneous information environment perform this aggregation by constructing a *representative agent* whose behavior replicates the aggregate behavior of the agents in the economy. We have not been able to do so in our environment of heterogeneous information. We approach this aggregation problem by an approximation based on the concept of bounded rationality. When contemplating his bid, a trader aggregates the bidding strategies in the market by taking a simple average of bidding strategies of players of different types. He uses this average strategy to decide his bid instead of a much more complex combination of strategies used by different types of participants. We motivate and defend this approach on the grounds of a *robustness criterion*. Using such a criterion, we show that this aggregation approach minimizes the maximum possible approximation error over all possible trading environments.

Third, our model produces predictions about the relationship between volume and expected price change that differ from the predictions of current market microstructure models. Most models, for example Kyle (1985) and Admati and Pfleiderer (1988) and (1989) predict a positive correlation between price change and volume in what is usually captured by the market depth parameter. The logic is that a market maker observing a large buy order will reason that there are information traders “camouflaged” behind this order and hence raise his ask price. Similarly, a large sell order will induce him to lower his bid price. We show that when one takes a different view in which liquidity traders
do not just submit their orders regardless of the price and when there is no designated market maker with inventory of the assets, the logic is entirely different. We show that volume is related to the change of the distribution around the average of traders' outlook on the market. If pessimists change their minds frequently to become optimists and then pessimists again, a large trade volume will be realized, independent of the price change. On the other hand, if some traders are "chronic" pessimists and others are "chronic" optimists, very small trading volume will be realized, also independent of the changes in price. These ideas are formally developed in section 6. Now, we briefly discuss the model and the results.

We study the trading environment of a purely speculative futures market with risk neutral traders. The futures contract is in zero net supply and the aggregate monetary gain is zero. Moreover, there are no motives for insurance and hedging. This is the simplest environment in which traders engage in the exchange of the futures contract solely because of differences in private information. It is well known that, in such a purely speculative market, if traders agree on the interpretation of the private signals, then no trade will occur. The common knowledge that trade is beneficial to all participants will lead each trader to conclude that his signal is not favorable and thus refrain from trade. This is the content of Milgrom-Stokey (1982) "no-trade" theorem.

We motivate trading in our model by differences among traders in the way they interpret information. Presented with the same piece of information, two different traders will reach two different conclusions about the probabilities of future events. Traders are interested in information about the value of the commodity underlying the contract at the time of maturity, which is a random variable we denote by $V$. During the lifetime of the contract, traders receive private information in the form of random variables, or signals, correlated with $V$. The main attribute of a trader in this world is how he links the signals to $V$. We capture this relation by the likelihood functions, which record the probability distribution of the signals conditional on $V$. Traders are differentiated on the basis of the likelihood functions, and these differences are the sole reason for trade.

The Economy is modeled as a continuum of traders, say $[0,1]$, who receive private information at a finite number of dates during the life of the contract which extends over the period $[0,T]$. Trading is conducted in a central exchange in which prices are
determined in a double auction. At any auction, the market can only accommodate a fraction $\Delta$ of the population and hence $\frac{1}{\Delta}$ auction are required to give each trader a chance to execute his order before the arrival of the next wave of information.

Traders are constrained to take net positions in only one contract. At a typical auction, long traders submit limit orders to sell two contracts at or above a limit price. On the other hand, short traders submit limit orders to buy two contracts at or below the limit price. Auctions are conducted when a balanced sample of orders of size $\Delta$ has accumulated on the floor, since there is no designated market maker. The market clearing price is thus the median of the quotes in the sample of limit orders, and the volume of trade is related to the measure of sellers (buyers) whose offers (bids) were below (above) the market clearing price. The price in the auction is made public to all traders, and the next group of orders is admitted to the market for the next auction. The process continues until all traders in the Economy have been given the chance to trade once in the Market. New information arrives to the Economy and the process is repeated.

A trader bases his bid on his private information and on the market clearing prices in previous auctions. A trader also conditions his bid on the information revealed to him by the fact that his order was executed. Thus, a trader who submits his order in ignorance of the market clearing price, avoids the “winner’s curse”. After learning the market price, each trader will have no reason to regret his current position.

The “no-regret” feature of a bidding strategy means that the buyer extracts information from the fact that the market price is below his bid. But market prices are not primitive source of information about $V$. They are endogenously determined, and to the extent that the trader understands how they were formed, he would be able to infer from them a statistic of the private signals embodied in the bids that ultimately led to the establishment of these particular prices. How would a trader extract this information? The question is easy to answer if there were only one type of traders in the market with understood bidding strategies. The median bid corresponds to the median private signal and by “inverting” the equilibrium strategy one could determine what the median signal is. When there are many types of traders with different bidding strategies, which strategy should one “invert” to infer the median of signals in the market? and why?

The above questions strike at the heart of the link between our formulation and what
one observes of the bahavior of a typical trader. A typical trader thinks of the Market as one entity that reacts to news in a way for which he has an intuitive feeling and as an independent personality with which he interacts. In a large anonymous market, a trader does not see himself as trading with Mr. X or Ms. Y. He deals, rather, with the Market. Observed behavior suggests that there is a "model" of a monolithic Market that traders use when acting strategically in determining their bids.

The issue extends beyond this intuitive feeling of the market. Even, if a trader knew exactly what the different bidding strategies are, inferring a statistic of private signals from the market price can be an exceedingly complex task. A small sample of traders of different types can have many configurations, and thus many possible mappings of signals to prices. Tabulating such mappings, assessing their relative likelihood of occurrence, and estimating a statistic of the private signals based on all possible configurations requires a great deal of computations. The number of such computations grows exponentially fast as the number of types and the sample size increase. This complexity calls for an approximate solution to this inference step. The introduction of this approximation is the essence of the bounded rationality approach that we advocate in this essay.

The approximate model that a trader uses to recover information from market prices should stand the test of time. If a trader used a model with success in particular circumstances only to discover that the model performs badly in a slightly different environment, he would have every reason to question its validity. A trader would prefer an approximate model that performs well over the largest possible set of circumstances. Alternatively, an approximation is preferred if it has the property that the magnitude of the maximum possible approximation error is minimized over all possible environments. We call this property "robustness" and we require that our approximation be robust.

A trader in our model aims to maximize his total expected profit at the end of the horizon when all the contracts are settled. A strategy for such a trader is a sequence of functions, one for each auction, that relate his bid to his private and any publicly available information. An equilibrium in any auction is a distribution of bidding strategies in the Economy, say $\theta^*$, such that if each trader maximizes his expected intertemporal payoff given $\theta^*$, the resulting distribution is again $\theta^*$. A perfect equilibrium is a sequence of bidding strategies that attain equilibrium in each auction.
We prove that a perfect equilibrium in this market exists under the assumption that the signals have the monotone likelihood ratio property (MLRP), that is higher signals imply higher expected value of the commodity at time $T$. This feature translates into monotonocity of the bidding functions, a property we use together with a mathematical result known as Tarski fixed point theorem to prove existence of an equilibrium.

Having established the existence of equilibrium, we proceed to analyze its properties. We note that different types of traders will interpret the endogenous parameters differently. Hence, we discuss the implications of the equilibrium from one fixed, yet arbitrary, point of view. All probabilistic statements will be made with reference to the chosen framework.

First, we note that the market clearing price is not fully revealing. The price reveals only the median of signals in a sample, which, in general, is not a sufficient statistic. Next, we consider the price sequence. Our "no-regret" proof for existence of equilibrium shows that a buyer today expects the price to go up tomorrow, whereas a seller today expects it to go down. Moreover, both of them might tomorrow change their positions and their views on the likely price movements. Therefore, in the case of heterogeneous traders, there does not seem to be a "consistent" way of characterizing price sequences.

We get, however, a sharp characterization of the sequence of trading volume between two information updates. We show that volume is closely linked to how private information is "distributed" in the Economy between two updates. In our market, where the traded asset is in zero net supply, and where there are no new participants, trade can only occur when an old seller decides to buy and simultaneously an old buyer decides to sell.

The old seller, who assumed his position earlier because he received information more pessimistic than average, will change his position only if he receives information more favorable than average, independently of the unconditional expected change in the value of the median of signals. This reversal of position is thus related to the reversals or fluctuations over time of a typical trader's private information around the average signal. We define this reversal property and show its relation to volume.

Finally, we show that volume and changes in prices are related to two different properties of the information structure. Volume is related to the reversal property, which is a
property of the sample path of realizations of the signals. On the other hand, changes in price and volatility of price are related to changes in the median of signals and volatility of median. There is no fundamental reason that these two properties should be linked. We construct an example in which expected volumes and expected price changes can be positively related, negatively related for all time periods, or have a relation that changes from positive to negative, or vice versa, over time. We see this result as an argument for the need for more explicit specification of the changes in the information distribution in models of price formation and volume analysis. We also think that this result suggests that all futures markets are not alike. The relationships we get for commodities futures might be different from that we get for futures contracts written on financial assets, for example.

This essay is organized as follows. Section 2 lays out the formulation of the model and sets the stage up for discussing information structures; a topic we deal with in section 3. In section 4, we introduce, motivate and defend the notion of bounded rationality. Armed with this notion, we move to section 5 to establish the existence of a perfect equilibrium in the model. In section 6, we discuss trading volume and the relations between volume and price change. Section 7 is reserved for concluding remarks.

In Appendix A, we review the existing theory of auctions and show that our environment does not fit with the existing formulations of double auctions on two counts. First, ours is a common value auction, whereas existing results treat private value models. Second, ours in an inherently asymmetric environment, whereas the literature deals primarily with symmetric auctions. We argue in this appendix for the need to analyze asymmetric double auctions. All proofs are contained in Appendix B.

2 Formulation

The model we consider is conceptually composed of two components: the economy and the market. We discuss each component and the interactions between them in the following.
2.1 The Economy

We study an economy over a time period \([0, T]\). The economy is populated by a large number of traders, which we model as points in the measure space \(([0, 1], \mathcal{B}[0, 1], \lambda)\), where \(\mathcal{B}[0, 1]\) is the Borel \(\sigma\)-field on \([0, 1]\) and \(\lambda\) is the Lebesgue measure. Traders take long or short positions in a futures contract that expires at time \(T\). The futures contract is written on a commodity whose value at time \(T\) is uncertain at any point \(t \in [0, T]\), and is denoted by \(V\). At a trading time \(t_i \in [0, T]\), which will be specified later, the futures contract specifies a futures price \(P_i\), at which the commodity will be exchanged for cash at time \(T\). We model the value of the underlying commodity \(V\) as a nonnegative random variable\(^1\) and we assume that \(V\) takes values in the interval \([\underline{V}, \overline{V}]\), where \(0 < \underline{V} < \overline{V} < \infty\) and hence \(\mathbb{E}[V] < \infty\). Agents in the economy are assumed to be risk neutral and are limited at any time to take positions in only one contract.

Information arrives to the economy at finite points in time, of total number \(N\), possibly randomly located on the interval \([0, T]\). At each information update epoch \(\{t_i; i = 1, 2, ..., N\}\), each agent receives a signal which is correlated with \(V\). After new information arrives to the economy, each agent is given the chance to trade once in the market.

2.2 The Market

Trading in the futures contract is conducted in a central exchange, in which prices and volumes are determined in a double auction. The market, however, can not accommodate all traders in the economy at once. Only a portion \(\Delta\), where \((0 < \Delta \leq 1)\), of all traders can be accommodated in the market at a time. We call \(\Delta\) the market capacity. The number of double auctions required to satisfy the orders of all traders in the period between information updates is \(1/\Delta\).\(^2\)

The demand functions for a trader \(\alpha \in [0, 1]\) are specified as follows. Let \(\mathcal{POS}_i(\alpha)\) for \(\alpha \in [0, 1]\) indicate the position of \(\alpha\) before the arrival of information at epoch \(t_{i+1}\). In other words, \(\mathcal{POS}_i(\alpha)\) is equal to \(+1\) if \(\alpha\) has a long position in the futures contract,

\(^1\)We implicitly introduce a probability space \((\Omega, \mathcal{F}, P)\) over which all random variables of the model are defined.

\(^2\)It is assumed here and in the rest of the exposition that \(1/\Delta\) is an integer \((L)\).
and is equal to \(-1\) if \(\alpha\) has a short position. At the auction that \(\alpha\) participates in, he quotes a price \(\pi_{\alpha}^{t+1}\), above which he is willing to sell \((1 - P\alpha S_{\alpha}(\alpha))\) contracts, and below which he is willing to buy \((1 - P\alpha S_{\alpha}(\alpha))\) contracts. For example, a trader \(\alpha\) who enters the auction with a short position in one contract can submit a limit order to buy two contracts at or below \(\pi_{\alpha}^{t+1}\) and change his position from short to long. If the market clearing price in the auction in which he participates turns out to be above his quoted price, he prefers to keep his original short position until new information arrives next period. We assume that \(P\alpha S_{\alpha}(\alpha) = 0\) for all \(\alpha \in [0, 1]\).

2.3 Market Dynamics

Information, both public and private, arrives to the economy in pulses at a finite number of dates in \([0, T]\). Each agent forms a reservation value of the futures contract based on his private information, the public information, and on his expectation of the behavior of agents in the rest of the economy. Agents, however, can only interact through the market; their reservation prices generate limit buy and sell orders, which are transformed by the market mechanism into market clearing prices and consummated trades.

Our model respects the physical fact that any market has a limited capacity for handling orders and that trading takes time. The time it takes any particular order to reach the trading floor and be considered in an auction is subject to different random factors. Therefore, at any market auction, the orders that determine the market price at that particular auction is a random sample of orders in the economy. We formalize this idea by introducing the following definition:

**Definition 1 (Market Sections)** A market section of size \(\Delta\), with \(0 \leq \Delta \leq 1\), is a (measurable) set of traders of measure \(\Delta\) selected at random from the \(\sigma\)-field \(B[0, 1]\).

After the arrival of new information at epoch \(t_i\), a random section of the economy of size \(\Delta\), denoted \(S_i^{(1)}\), is selected to trade in the first auction, denoted \(A_i^{(1)}\). After the orders of traders in \(S_i^{(1)}\) are processed, these traders are not allowed to enter the market until the next new information update epoch. The market price determined in the first auction is declared and a second (random) section of the economy, \(S_i^{(2)}\), is admitted to the market.
for bidding in auction $A_{i}^{(2)}$. This process continues until all traders in the economy have been given the chance to submit their bids.

We make the following assumptions about the structure of the market.

**Assumption 1 (Balanced Auctions)** At each double auction, the measure of the (potential) buyers is equal to the measure of the (potential) sellers, which is equal to $\Delta/2$.

We note that in our setup, at any time $t_{i}$, before trading, the measure of traders with long positions is equal to the measure of traders with short positions, which is $\frac{1}{2}$, since the futures contract is of zero net supply. Assumption 1 only implies that the random sampling procedure which takes orders from the Economy to the Market does not create any temporary imbalance in the auctions.

This assumption is motivated by the fact that in actual futures trading pits there is no designated market makers. The temporary imbalance in orders on the trading floor is absorbed by the so-called scalpers. Scalpers are floor traders who stand ready to take positions against incoming orders. Scalpers, however, carry no inventory and liquidate their positions very quickly. We ignore the role of the very short term market makers in our analysis.

**Proposition 1 (Market Clearing Price)** Using assumption 1, the market clearing price in each auction is the median of quotes (buy and sell limit prices) in the auction.

**Proposition 2 (Trade Volume in an auction)** For trader $\alpha \in [0,1]$, let $B\mathcal{P}OS_{i}(\alpha)$ indicate the position of $\alpha$’s quote relative to the median of quotes in the auction in which he participates after information arrival at time $t_{i}$. In other words, $B\mathcal{P}OS_{i}(\alpha)$ is equal to $+1$ if $\alpha$’s quote is above the median, and is equal to $-1$ if $\alpha$’s quote is below the median. The volume of trade at any auction $A_{i}^{(j)}$ between $t_{i}$ and $t_{i+1}$, denoted $VOL_{i}^{(j)}$, is given by:

$$VOL_{i}^{(j)} = \lambda\left(\{\alpha \in S_{i}^{(j)}: |B\mathcal{P}OS_{i-1}(\alpha) - B\mathcal{P}OS_{i}(\alpha)| \neq 0\}\right).$$

We also make the following assumption about market dynamics:

**Assumption 2 (Enough Absorption Time)** We assume that the total trading time required to conduct $L$ auctions is smaller than the minimum interarrival time between consecutive information pulses.

21
Assumption 2 says that the market is of such a physical design that it gives all traders the chance to transact on the basis of information before their reservation prices are affected by the arrival of new information.

2.4 Why Trade?

It is very important to understand why traders participate in such a market. As described above, this market is a purely speculative market since the aggregate monetary gain is zero and since there is no insurance motives. If traders have, in the terminology of Milgrom and Stokey (1982), concordant beliefs, then no trade will occur. Traders are said to have concordant beliefs, if when they all receive a signal which is correlated with the value of the underlying commodity $V$, they interpret this signal in the same way. In the futures market described above, the first round of trade will lead to a Pareto optimal allocation of contracts. With concordant beliefs further trades on the basis of private information are side bets between traders. Risk neutral traders will be indifferent to such bets and a trivial equilibrium with no trade will obtain.

Models of futures markets avoid this difficulty in two ways. Some models introduce noise traders, who are traders with unmodeled motives for trade. Their effect in a model only appears in their random orders. This approach does not, in my opinion, avoid the no trade theorem, it just pushes the question of trade motives outside the formal framework of the given model. Examples of such approach include Kyle(1985), and Miller and Grossman (1988). The informal motivations for noise trader participation in the market are related to liquidity shocks and different consumption life cycles among traders. It seems to me that such arguments are suitable for motivating trades in an investment asset like common stock. Futures contracts are relatively short-lived and are not generally held for investment purposes.

The second approach motivates trades by differences in attitudes toward risk and by differences in endowments. A group of risk averse traders are endowed with the underlying commodity, and members of this group use the futures contract to buy some form of insurance. Another group of traders provide this insurance in exchange for a positive premium. An example of such models is Bray(1981).

This second approach deals with futures markets as mechanisms for intertemporal
allocation of risks. The purpose of this work is to study specifically differences in information and their effects on prices and volume, I would like to isolate the informational phenomena from the hedging/risk sharing phenomena. Later, after understanding the effects of information differences in isolation, I hope that I would be able to enrich the analysis by adding more motives for trade.

For now, to focus on the pure informational phenomena, and to avoid conflict with the Milgrom-Stokey (1982) no-trade theorem, I motivate trading in the market by differences in beliefs and information structures. These are the issues we discuss now.

3 Information Structures

Information in our model is of two kinds: exogenous and endogenous. Exogenous information includes both private and public signals, whereas endogenous information is information about market prices and volumes. The basic source of uncertainty in the model is the value $V$ of the underlying commodity, which will be revealed at time $T$. In the meantime, an agent $\alpha \in [0,1]$ can obtain information about $V$ from a variety of sources; weather reports, supply or demand news, or consultants’ opinions. In addition, he has access to public news about factors that affect $V$. In general, one could represent trader $\alpha$’s information acquisition by the realization of a stochastic process which takes values in an abstract measure space. To simplify our analysis, we make the following assumption.

**Assumption 3 (Finite Number of Information Updates)** Agents can obtain new information only at a finite number of points in time, $\{t_i \in [0,T], i = 1,2, \cdots, N\}$. Furthermore, we assume that at each $t_i$, agents simultaneously receive their new private signals and possibly any new public information.

We model public information at each time $t_i$ by a real-valued random variable $Y_i$, which is correlated with $V$. We model private information at time $t_i$ by a family of real-valued random variables $\{X^\alpha_i, \alpha \in [0,1]\}$ such that:

1. For all $\alpha \in [0,1]$, $X^\alpha_i$ has a conditional distribution $F_i(. \mid V)$, and
2. Regardless of the identity of the traders who observe particular realizations of members of the family $X_i^\alpha$, we have

$$\lambda\left(\{\alpha \in [0, 1]: X_i^\alpha \leq \xi\} \mid V\right) = F_i(\xi \mid V) \quad (2)$$

Remark 1 Using a real valued process to model information flow is based on the assumption that at any time, any two signals are “comparable” as defined by Milgrom (1981a).

Remark 2 This approach to modeling the private signals of agents implies that at any information update epoch $t_i$, there is no aggregate uncertainty in the economy. In other words, all the information required to determine the value of $V$ is available in the economy. The economy as a whole has complete information about $V$, but each individual trader has only incomplete information. If hypothetically, each trader were to announce his signal publicly, then the distribution of signals would become public knowledge, and the value of $V$ would be instantly revealed. We view the market, in this essay, as a mechanism for obtaining a random section of signals (embodied in bids via the equilibrium strategies) at a time, and reducing the information in these signals to a few statistics; prices and volumes.

Remark 3 The internal consistency of these conditions and their implications for the dependence of signals of different agents is discussed in Appendix C.

We assume that each trader $\alpha \in [0, 1]$ is endowed with the ability to observe one realization of the random variable $X_i^\alpha$ at each time $t_i$. Hence, the temporal revelation of information to a trader $\alpha$ is given by the two discrete time stochastic processes: $X^\alpha$ with components $[X_1^\alpha, X_2^\alpha, \cdots, X_N^\alpha]$, and $Y$ with components $[Y_1, Y_2, \cdots, Y_N]$. Traders, however, differ in the way they interpret these observations. To formalize this notion, we make the following definitions:

Definition 2 (Information Structure) An information structure for individual $\alpha \in [0, 1]$ is the conditional distribution of $X^\alpha$ and $Y$ given $V$, denoted by $F^\alpha(X^\alpha, Y \mid V)$. It is assumed that for all $\alpha \in [0, 1]$, $F^\alpha(X^\alpha, Y \mid V)$ has a positive density which we denote by $f^\alpha(X^\alpha, Y \mid V)$. 24
Definition 3 (Distribution of Information Structures in the Economy) Let $\Gamma$ be a finite set of information structures, i.e. $\Gamma$ is a finite set of conditional distribution functions of $\mathcal{X}^\alpha$ and $\mathcal{Y}$ given $V$. Let $\Sigma(\Gamma)$ be the discrete $\sigma$-field on $\Gamma$ (The $\sigma$-field generated by singletons). A distribution of information structures in the economy is a (measurable) mapping:

$$\Psi: ([0, 1], \mathcal{B}[0, 1], \lambda) \rightarrow (\Gamma, \Sigma(\Gamma)).$$

We interpret $\Psi(\alpha)$ to be the information structure with which individual $\alpha$ is endowed and we call it his type. Note that the mapping $\Psi$ induces a (probability) measure $\mu$ on $\Sigma(\Gamma)$. For $\gamma \in \Gamma$, $\mu(\gamma)$ is the measure of traders with information structure $\gamma$. Since $\Gamma$ is finite, we can enumerate this measure in $(\mu_1, \mu_2, \cdots, \mu_k)$ corresponding to $(\gamma_1, \gamma_2, \cdots, \gamma_k)$, where $k \geq 2$.

For $\alpha$ and $\beta \in [0, 1]$, with $F^\alpha(\mathcal{X}^\alpha, \mathcal{Y})$ different from $F^\beta(\mathcal{X}^\beta, \mathcal{Y})$, we could interpret this difference as a difference in the way $\alpha$ and $\beta$ interpret and analyze the signal $(\mathcal{X}, \mathcal{Y})$. If $\alpha$ and $\beta$ had the same prior on $V$, and if the realization of $\mathcal{X}^\alpha$ were equal to the realization of $\mathcal{X}^\beta$, then they would interpret their signals differently and thus end up having different posterior assessment of the expected value of $V$.

Another way of interpreting this difference is assuming that $\alpha$ and $\beta$ have different technologies of obtaining their signals $\mathcal{X}$. For each realization of $V$, these technologies produce the signals $\mathcal{X}$ with different likelihoods for $\alpha$ and $\beta$. In the rest of the essay, we will suppress the dependence of the likelihood functions on the identity of the agents; they will be distinguished on the basis of the agent’s type only. We will use the notation $F^\gamma(\cdot | V)$ to denote the likelihood functions for an agent whose information structure is $\gamma_i$, for $i = 1, 2, \cdots, k$.

Finally, we make the following assumptions about the nature of the information structures.

Assumption 4 (Temporal Sufficiency) For all agents $\alpha \in [0, 1]$, and for all types of information structures $\gamma_i, i = 1, 2, \cdots, k$, the random variables $X_i^\alpha$ and $Y_j$ are sufficient statistics for the family of conditional probability density functions:

$$f^n(X_j^\alpha, X_{j-1}^\alpha, \cdots, X_1^\alpha, Y_j, Y_{j-1}, \cdots, Y_1 | V).$$  \hspace{1cm} (3)
Assumption 5 (Monotone likelihood Ratio Property) For all $\gamma_i, i = 1, \cdots, k$, $F_j^n(X_j \mid V)$ has the strict monotone likelihood ratio property (MLRP). In other words, for all $\gamma_i$, $f_j^n(x_j \mid v)/f_j^n(x_j \mid v')$ is strictly decreasing in $x_j$ for $v' > v$, and strictly increasing in $x_j$ for $v > v'$.

Remark 4 (First Order Stochastic Dominance Property) Assumption (5) implies that for all $\gamma_i, i = 1, \cdots, k$, $F_j^n(X_j \mid V)$ has the first order stochastic dominance property. In other words, for all $\gamma_i$, $F_j^n(x_j \mid v) \leq F_j^n(x_j \mid v')$ for $v' > v$. See Milgrom (1981a).

Assumption 4 implies that for any agent $\alpha \in [0, 1]$, at any time $t_i$, statistical inferences about $V$ and about future development of market prices depend on public and private information gathered over time only through the value of the most current informational variable (signal). One could accept this assumption as a modeling choice in which a signal at time $t_i$ is redefined so as the condition of temporal sufficiency is satisfied. As mentioned earlier, when one models information at any point in time $t_i$ by a real-valued random variable, $X_i$, one already assumes that $X_i$ is a sufficient statistic for all information in other forms that an agent might have. Accepting such an assumption of categorical sufficiency of $X_i$, makes it plausible to accept its temporal sufficiency as well.

One could also follow the lead of Milgrom and Weber (1982) in interpreting the informational variable $X_i$ as a value estimate, which is the only piece of information available to any agent. This view makes assumption 4 a natural condition to impose. Finally, one could also imagine that the informational variable $X_i$ is short-lived. The new information that arrives at $X_{i+1}$ makes the value of $X_i$ useless for the purpose of statistical inference. In such an environment, assumption 4 is also a natural condition.

Assumption 5 simply says that at any information update epoch $t_i$, and under any interpretation of the signals, higher values of both private and public signals imply higher expected values of $V$ at $T$. In other words, agents of all types agree that higher values of $X_i^*$ and $Y_i$ are suggestive of higher values of $V$, but they differ in their expectations of $V$. Again, this is the sole reason for trade in this model.
4 Bounded Rationality and Market Averaging

4.1 The Case for Bounded Rationality

Recall the structure of our model. At each information update epoch $t_i$, agents receive signals $X_i^x$ about $V$. Based on these signals, together with public information and past prices, agents submit their bids, and in each auction $A_t^{(j)}$ the market price is the median of a section of bids of size $\Delta$. Thus the price reflects the information of the traders. Now consider an agent in this model, who is about to participate in auction $A_t^{(j)}$, and assume that this agent were told that the market clearing price in this auction will be $p$; what inferences could this agent make about $V$?

The agent understands that $p$ is the median of bids in a section of size $\Delta$. However, in order to estimate the expected value of $V$ conditional on the price, the agents needs to know the realization of the median of a $\Delta$-section of signals, and not the median of a $\Delta$-sections of bids. In our model signals are the primitive sources of information about $V$. The bids are just reflections of the signals in a manner determined by the equilibrium bidding strategies. In the case of a market with traders of the same type, recovering the median signal from the median bid is relatively easy. Since in a symmetric equilibrium all agents have the same trading strategy, and assuming that the equilibrium strategy is increasing in the private signal $^3$ the median signal can be obtained by inverting the equilibrium strategy$^4$. The agent then uses this value, together with his own signal, to form an expectation of $V$ and to determine his optimal bid.

The situation is very different when traders have different information structures; the condition necessary to sustain trade in our model. In a market with $k$ types of traders with measures $(\mu_1, \mu_2, \ldots, \mu_k)$, inferring the median signal from the median bid is much more difficult. To be sure, the median bid which determined the price could have been submitted by an agent of any type. Hence the signal corresponding to this particular bid has an expected value of $x^* = \sum_{i=1}^{k} \mu_i \pi_{i}^{-1}(p)$, where $\pi_i$ is the equilibrium bidding strategy for agents of type $i$.$^5$ The value $x^*$ need not, in general, be the median of the signals received by the agents whose bids are included in the market section. Since the

---

$^3$This condition obtains in equilibrium.

$^4$This inverse is well-defined for all values of the publicly known information; $Y_t$ and prices.

$^5$Here we drop the auction indices and the public information variables for ease of notation.
equilibrium strategies differ from one type to another, it is possible for an agent who receives a signal much lower (or higher) than the median signal to submit a bid that becomes the median of the section of bids. An agent trying to recover the median of the sample of signals should take such possibilities into account.

The number of possibilities in which the ordering of bids does not correspond to the ordering of the corresponding signals increases very quickly as the number of types \( k \) increases. To see the problem clearly, assume here that the market section is actually composed of a finite number \( n \) of subsections, each of which can be of one homogeneous type. Such a sample can be configured in \( k^n \) possible ways with \( k \) possible types. Moreover, in each configuration, there is a strictly positive probability that the order of signals would not be preserved when signals are transformed to bids via the equilibrium bidding strategies.

If one measures the complexity of the inference problem facing an agent in the market by the number of the possible configurations corresponding to a fixed sample of \( n \) possible subsections, or by \( k^n \), one realizes that a market with a few types is orders of magnitude more complex than a homogeneous market. The diversity of types in the economy that brings more incentives and motives for trade brings also inherent complexity to the model.

Faced with the complexity of the inference step, a trader might use an approximate "market model" to guide his bidding behavior. By introducing such an approximation, a trader greatly reduces the amount of reasoning required to reach an optimal decision. We would like to model such approximations and establish meaningful criteria for accepting them. This is the program we consider next.

4.2 The Approach

We model trading in the futures market described before as an anonymous sequential double auction. For each \( \alpha \in [0, 1] \), and at each auction \( A_i^{(j)} \) (see section 2.3), an auction strategy \( \pi_\alpha^{(i,j)} \) is a (measurable) mapping:

\[
\pi_\alpha^{(i,j)}: (\mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{i-1}, \mathcal{B}(\mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{i-1}), \lambda') \to (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)),
\]

where \( \mathcal{B}(\mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{i-1}) \) is the Borel \( \sigma \)-field on the product space \( \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{i-1} \), and \( \lambda' \) is the Lebesgue measure on that space.
The auction strategy \( \pi^{(i,j)}(x, y, p^{(1)}, \ldots, p^{(j-1)}) \) is the bid that \( \alpha \) quotes in the market when his private signal at \( t_i \) is \( X_i^\alpha = x_i \), the public information at \( t_i \) is \( Y_i = y_i \), and when the prices in the sequence of auctions starting at \( t_i \) are \([p_i^{(1)}, p_i^{(2)}, \ldots, p_i^{(j-1)}] \).\(^6\) For each agent \( \alpha \in [0, 1] \), the collection of auction strategies \( \pi^{(i,j)}_\alpha \), for \( i = 1, 2, \ldots, N \) and \( j = 1, 2, \ldots, L \), is called the agent’s strategy and is denoted by \( \pi_\alpha \).

Denote the space of strategies by \( \Pi \), endow \( \Pi \) with a measure structure\(^7\), and let the space of probability measures on \( \Pi \) be \( \Theta(\Pi) \). A choice of strategy by each \( \alpha \in [0, 1] \) induces a certain probability distribution \( \theta \in \Theta(\Pi) \). An equilibrium of the anonymous sequential double auction game is a distribution of strategies \( \theta^* \in \Theta(\Pi) \), with the property that if each trader maximizes his intertemporal expected payoff given \( \theta^* \), the resulting distribution of strategies is again \( \theta^* \).

The anonymous sequential double auction is an asymmetric bidding game, since one expects traders with different information structures to have different bidding strategies. We will be concerned with equilibria in which traders endowed with the same information structure adopt the same bidding strategy. Our approach to the analysis of this game is based on the following two concepts:

**Definition 4 (Average Bidding strategy)** In the anonymous sequential double auction game, let there be \( k \) types of traders with measures \((\mu_1, \mu_2, \ldots, \mu_k)\). Let the equilibrium bidding strategy for type \( l \) be \( \pi_l \). We define the average (market) equilibrium bidding strategy \( \pi_A \) as:

\[
\pi_A^{(i,j)}(x, y, p^{(1)}, \ldots, p^{(j-1)}) = \frac{\sum_{l=1}^{l=k} \left[ \pi_l^{(i,j)}(x, y, p^{(1)}, \ldots, p^{(j-1)}) \ast \mu_l^2 \right]}{\sum_{l=1}^{l=k} \mu_l^2} \tag{4}
\]

Note that for every realization \( v \) of \( V \), and at any auction \( A_i^{(j)} \), the distribution of \( X_i^\alpha \) and \( Y_i \) conditional on \( v \) induces a distribution of the bids that traders of different types quote in \( A_i^{(j)} \). It also induces a distribution on the bids drawn from the fictitious bidding rule \( \pi_A^{(i,j)} \).

**Assumption 6 (Bounded Rationality)** Consider the anonymous sequential double auction game. Let bidders of type \( l \), where \( l = 1, 2, \ldots, k \), adopt the bidding strategy \( \pi_l \). Let

\(^6\)The reader is reminded that each trader gets to bid only once between \( t_i \) and \( t_{i+1} \).

\(^7\)The details of a measure structure on \( \Pi \) are not important for our analysis.
the average (market) bidding strategy be \( \pi_A \). Let \( S_i^{(j)} \) be a market random section of bids of measure \( \Delta \) tendered in auction \( A^{(j)}_i \). Let \( FS_i^{(j)} \) be a random section drawn from the (fictitious) distribution of bids implied by \( \pi_A^{(i,j)} \). We make the bounded rationality assumption that when agents make inferences they substitute the distribution of any statistic of \( S_i^{(j)} \) by the distribution of the same statistic of \( FS_i^{(j)} \). In particular, we assume that agents make the approximation:

The distribution of the median of \( S_i^{(j)} \) is equivalent to the distribution of the median of \( FS_i^{(j)} \), for all \( i = 1, 2, \ldots, N \) and \( j = 1, 2, \ldots, L \).

4.3 Robustness

Let us recall how agents are supposed to behave in a full rationality world. In a typical model, one specifies a set of exogenous variables, in our case these are the likelihood functions (information structures), utility functions and prior beliefs. Given these exogenous parameters, agents take actions that lead to the determination of endogenous variables. In our model, agents maximize their expected profits conditional on their signals and on the fact that they are buyers or sellers, and submit bids which determine the price. In a full rationality model, agents have an understanding of the relationship between exogenous and endogenous variables. An important implication of full rationality is that the understanding of agents is confirmed in equilibrium. In other words, the relationships that emerge from the model are what the agents understand them to be.

One could depart from this full rationality framework in many directions. In a learning model, agents discover that their prior understanding of the relationships between endogenous and exogenous variables is not confirmed, and they modify their understanding gradually. One examines the limiting result of this learning behavior, see for example Bray and Kreps (1987). Another departure could be motivated by the fact that the relationships of the model are very complex, and even if agents understand the structural nature of this complexity, they can not cope analytically with it, because of inherent intractability. A case in point is our model. In such circumstances, it seems plausible that agents would approximate these complex relationships by simplified ones that have the same essence, but with a manageable analytical structure.

If one accepts approximations within a model as a valid analysis approach, then one
accepts the possibility of a loss of consistency in the model. By the very nature of approximation, agents' approximate understanding of the structure of the endogenous variables in a model is not exactly what this structure turns out to be in the same model. Since in the class of models we consider; models of asymmetric information, the endogenous variables we study are only probabilistically determined, one is led to compare, on the basis of limited observations, various statistical representations of these variables. The next logical step is, then, the establishment of reasonable criteria for accepting models formulated along this approximation approach. I propose the following Robustness Criterion.

The robustness criterion can be motivated by the observation that in the bounded rationality approach, an approximation is introduced before the model is completely solved, and hence the approximation is independent of the specification of the exogenous environment. Recall our model. Agents construct the market response to a signal \( x \), or the average bidding strategy, as an average of the responses of traders of different types to \( x \). This averaging is independent of the information structure and the underlying value \( V \). Now, if agents were not completely sure of the specification of the exogenous environment, would they adopt the same approximate view? or would they adopt another view that is completely different? We would like our agents to adopt an approximate view that is good for the largest possible set of specifications of the exogenous variables. Equivalently, we would like an approximation that is least sensitive to modeling errors. An approximation with this property is said to be a robust approximation.

To apply the robustness criterion to our model, we need to understand the nature of the approximation we are introducing. Fix an auction \( A_i^{(j)} \). For any value \( V \), and from the point of view of any trader, there is a distribution of private signals in the market given by, say, \( F(x \mid V) \). Different types of traders will respond differently to the same signal \( x \). Hence bids \((b)\) from agents of type \( l \), which is of measure \( \mu_l \), and equilibrium strategy \( \pi_l(x) \), will be distributed by \( G_l(b \mid V) \), where \( G_l(b \mid V) \) is conditional distribution function of bids induced from \( F(x \mid V) \) by the transformation \( \pi_l(x) \).

The distribution of bids in the market, for any \( V \), is a mixture of \( k \) distribution with weights \([\mu_1, \mu_2, \ldots, \mu_k]\), and corresponding distribution functions \([G_1(b \mid V), G_2(b \mid V), \ldots, G_k(b \mid V)]\).

---

\( ^8 \)Remember that \( F(\cdot \mid V) \) is different for different types.
A section of bids of size $\Delta$ has a median which is distributed around a mean equal to the median of the mixed distribution and which has a certain variance. Note that since each $\pi$ is a strictly increasing function of the private signal $x^9$, the conditional median of each component distribution $i$ is $\pi_i(x^*)$, where $x^*$ is the median of the distribution of signals conditional on $V$. However, the median of the mixture of these distributions is not a function of only the median of the individual distributions. In fact, its location depends also on the distribution functions $G_i(b \mid V)$.

We would like to construct an estimator of the median of the mixed distribution given only knowledge about the locations of the component distributions, and absent any specifications of the component distribution functions. Our approximation thus implies that the distribution of the median of a section of bids has a mean equal to the approximate (estimated) median of the mixed distribution. This is a source of inconsistency since the approximate median is not, in general, equal to the true median. The approximate median does not take into account the structure of the component distribution functions. Therefore, we would like the chosen approximation to be robust in the sense that it minimizes the maximum possible error over all the specifications of the distribution functions. This is a necessary requirement since the individual distribution functions are endogenously determined and can not be used before completely solving the model.

The robustness criterion applied to our model boils down to the following proposition:

**Proposition 3 (Robust Approximation of A Median)** Let $G_1(z), G_2(z), \ldots, G_k(z)$ be distribution functions of the random variable $Z$, with strictly positive density functions $g_1(z), g_2(z), \ldots, g_k(z)$, and with median $z_1^*, z_2^*, \ldots, z_k^*$, respectively. Let $G(z)$ be a mixture of $G_i, i = 1, \ldots, k$, with weights $\mu_1, \mu_2, \ldots, \mu_k$. Let $\mathcal{E}$ be the set of estimators of the median of $G$, which are functions only of $z_1^*, z_2^*, \ldots, z_k^*$. The robust element of $\mathcal{E}$ that minimizes the maximum possible value of the approximation error over the set of all $k$-tuples of distribution functions $G_i(.)$ is given by:

$$z^* = \frac{\sum_{i=1}^{k} \mu_i^2 z_i^*}{\sum_{i=1}^{k} \mu_i^2}.$$

**Proof.** See Appendix B  

---

9 This fact is proved in a later section.
5 Equilibrium

5.1 Individual’s Problem

In a bidding game, a trader’s bid has two effects. First, a bid determines the probability of winning an auctioned item in a single auction, whereas in our formulation of a double auction of an asset of zero net supply, a bid determines the position of a trader, either long or short. Second, a bid has an effect on the price at which a bidder ultimately transacts. This effect depends on the pricing rule in the particular bidding game. The interplay between these two effects is at the core of the equilibrium analysis of auctions.

In keeping with the spirit of competitive markets, the bid of an individual trader in an anonymous market with a continuum of traders should not affect the equilibrium price. The random selection mechanism of market sections implements this intuition. Since any market section is of a finite measure $\Delta$, no one trader, of measure zero, can affect the value of the median of bids, and hence the market clearing price. Thus an individual price taker in this sense is only left with the strategic impact of his bid on his net position.

Market clearing price in each auction conveys information to traders. In any particular auction, however, the participating traders must submit their bids in ignorance of the price at which they will ultimately transact. As Milgrom (1981b) points out, this feature distinguishes the bidding game from Rational Expectations Equilibrium models. Nevertheless, in equilibrium, traders conditions their bids on the event that their signal is higher (lower) than the median signal among the participants in the auction. We will show that a trader who takes such an event into account would have no incentive to revise his bid after learning the actual price. This “no regret” feature of the equilibrium bidding strategy means that a trader acts as if he knew in advance the transaction price. Hence, we can interpret the bids as either limit orders or as market orders after observing a price which is not expected to change before the order is executed.

As mentioned earlier, traders are constrained to have a net position in only one futures contract at any point in time. This implies that at some auctions, some traders will submit bids that affect the market clearing price, but who will end up not trading in these auctions. A trader with a long position before an auction, whose bid turns out to be above the median bids will not be allocated a new unit of the contract. His bid has
contributed to determining the price but he did not transact in the market.

To see the above mentioned point more clearly, imagine that market clearing in any market section as taking place over a finite interval. A small number of traders in the market section have access to the floor, where they transact and set a price which is a representative of the median valuation of all potential traders in this particular market section. This price is then broadcast to all members of the market section over trading screens. Other traders in the section then compare this price to their reservation prices, as if they submit silent bids, and then decide whether to transact in the market or wait.

If a trader has a long position and if he thinks that the screen price is too high relative to his reservation price, he would then decide to submit a market sell order of two units at a price very close to the screen price. If, on the other hand, he thinks that the screen price is too low, and since he is constrained to hold only one unit of the contract, he will decide to hold on to his position until the prices move up above his reservation price. This comparison between the screen price and the reservation price of any trader is formalized in our model by the structure of the double auction in which all traders in a market section participate. We will show that after the trader learns of the market clearing price in his auction, he will not revise his valuation of the value of the contract. This is equivalent to a trader first learning the market price, and then deciding whether to change his position or not at this particular price.

Recall that $\mathcal{POS}_i(\alpha)$ is trader’s $\alpha$ position during the no-new-information interval $[t_i, t_{i+1}]$. An individual trader’s problem is thus to choose a sequence of auction bidding strategies $\pi^{(i,j)}_\alpha$, denoted by $\pi^*_\alpha$, to solve:

$$\max_{\pi} \mathbb{E} \left[ \sum_{i=1}^{i=N} [(p_{i+1} - p^{(i)}_i)\mathcal{POS}_i]\right],$$

where $p_i^{(j)}$ is the market clearing price at the auction in which $\alpha$ happens to participate, $p_{i+1}$ is the expected price in the auctions during the interval $[t_{i+1}, t_{i+2}]$, and where the expectation is taken using $\alpha$’s information structure. Each auction strategy depends on the prices in the sequence of auctions up to the current one, and on public and private information received up to date. The assumption that new information makes older information obsolete, however, implies that a trader’s bid will depend only on the prices in the auctions starting after the most recent update and on the most current
information. The usefulness of historic prices and trends is ruled out by assumption. Moreover, the assumption of a continuum of traders allows us to simplify a trader's intertemporal problem as follows.

**Proposition 4** The multiperiod decision problem in 5 is equivalent to a sequence of single period problems given by:

$$
\max_{\pi^{(i,j)}} \mathbb{E} \left[ \left( p_{i+1} - p_{i}^{(j)} \right) \mathcal{P} \Omega_{i} \mid \mathcal{F}_{\alpha}^{(i,j)} \right],
$$

where $\mathcal{F}_{\alpha}^{(i,j)}$ is all public and private information that individual $\alpha$ has before the start of auction $A^{(j)}_i$.

**Proof.** See Appendix B

An equilibrium in this sequential double auction game is a Bayesian-Nash equilibrium in each auction with the additional property that the equilibrium strategy will be followed at any time after each possible history of the game. We provide a formal definition.

**Definition 5 (Perfect Equilibrium)** Let $\Pi$ be the space of bidding strategies, endowed with a metric structure, let $\mathcal{B}(\Pi)$ be the Borel $\sigma$-field on $\Pi$, and let $\Xi(\Pi)$ be the space of probability measures on $(\Pi, \mathcal{B}(\Pi))$. Let $u^{(i,j)}(\pi, \xi)$ denote the expected payoff to player $\alpha$, conditional on all information available prior to the start of auction $A^{(j)}_i$, when he adopts the strategy $\pi$, and when the distribution of bidding strategies in the market is $\xi$. A distribution of strategies $\xi^* \in \Xi(\Pi)$ is a perfect equilibrium in the anonymous sequential double auction game if there exists a (measurable) mapping:

$$
\Psi: ([0, 1], \mathcal{B}[0, 1], \lambda) \rightarrow (\Pi, \mathcal{B}(\Pi))
$$

such that:

1. For all $\alpha \in [0, 1]$, and for all auctions $A^{(j)}_i$, $u_{\alpha}(\Psi(\alpha), \xi^*) \geq u_{\alpha}(\pi', \xi^*)$, for $\pi' \in \Pi$.
2. $\lambda \circ \Psi^{-1} = \xi^*$. 

35
5.2 Existence of Perfect Equilibrium

In this section we demonstrate the existence of a perfect equilibrium for the sequential anonymous double auction game. We show that the reservation price of each trader is a nondecreasing function of his private signal. This characterization is based on the assumption that the signals have the monotone likelihood ratio property and on the bounded rationality assumption. We also show that this characterization is preserved under the random selection mechanism which admits traders to the market for bidding, and hence the market bid of a trader is also a nondecreasing function of his private signal. We utilize this characterization, together with a mathematical result known as Tarski fixed point theorem, to prove existence of an equilibrium. First, we provide some facts that will be useful later.

**Lemma 1** Let $X$ be a random variable whose conditional distribution function and conditional density function are $F(X \mid V)$ and $f(X \mid V)$, respectively. If $X$ has the MLRP, then the function $\varphi(x) = \frac{F(x \mid v) - c_1}{F(x \mid v') - c_2}$, for $c_1, c_2 \in [0, 1]$, $c_2 \geq c_1$, is nonincreasing in $x$ for $v' > v$, and for all $x$ such that $F(x \mid v) \geq c_1$ and $F(x \mid v') > c_2$.

**Proof.** See Appendix B

**Proposition 5 (Conditional MLRP)** Let $S_i^{(j)}$ be the market section number $j$, after the arrival of new information at $t_i$, and let $M_i^{(j)}$ be the median of signals received by individuals in $S_i^{(j)}$. For all information structures $\gamma_l \in \Gamma, l = 1, \cdots, k$, under assumption $5$, and conditional on the values of $M_i^{(1)}, \cdots, M_i^{(j-1)}$, the random variable $M_i^{(j)}$ has the MLRP.

**Proof.** See Appendix B

Finally, we introduce the concepts of a valuation process and a fundamental valuation process.

**Definition 6 (Valuation Process)** For every auction $A_i^{(j)}$, for all individuals $\alpha \in [0, 1]$, and for every increasing function $\varphi: \mathbb{R} \to \mathbb{R}$, define the valuation process:

$$VAL^{(i,j)}_{\alpha} = E[\varphi(V) \mid F^{(i,j)}, M_i^{(j)}],$$

(7)
where $\mathcal{F}_\alpha^{(i,j)}$ is all public and private information that individual $\alpha$ has before the start of auction $A_i^{(j)}$, and where $M_i^{(j)}$ is the median of signals in section $S_i^{(j)}$. By virtue of assumption 4, we can write:

$$\mathcal{VAL}_\alpha^{(i,j)} = \mathbb{E}[\varphi(V) \mid X_i^\alpha, Y_i, p_i^{(1)}, p_i^{(2)}, \ldots, p_i^{(j-1)}, M_i^{(j)}],$$

(8)

where $p_i^{(k)}$ is the market clearing price in auction $A_i^{(k)}$.

Definition 7 (Fundamental Valuation Process) For every auction $A_i^{(j)}$, for all individuals $\alpha \in [0,1]$, and for every increasing function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$, define the fundamental valuation process:

$$\mathcal{FVAL}_\alpha^{(i,j)} = \mathbb{E}[\varphi(V) \mid X_i^\alpha, Y_i, M_i^{(1)}, \ldots, M_i^{(j-1)}, M_i^{(j)}],$$

which we write as $\mathcal{FVAL}_\alpha^{(i,j)}(X_i^\alpha, Y_i, M_i^{(1)}, \ldots, M_i^{(j)})$.

The fundamental valuation process is the expected value of $\varphi(V)$ that individual $\alpha$ estimates if he knows the realizations of the medians of signals in the different market sections, including the median of signals of individuals in the auction yet to begin. This function will play an important role in the following analysis. The fundamental valuation process has a crucial monotonicity property that we record in the following proposition:

Proposition 6 (Monotonicity of Fundamental Valuation Process) For all information structures, and for all auctions, the fundamental valuation process $\mathcal{FVAL}_\alpha^{(i,j)}(X_i^\alpha, Y_i, M_i^{(1)}, \ldots, M_i^{(j)})$ is increasing in all its arguments, and strictly increasing in $X_i^\alpha$.

PROOF. The proof follows immediately from proposition 5, the independence of the random selection process and the variables $X_i^\alpha$ and $Y_i$, and from theorem (3.2) in Milgrom (1981b). $\blacksquare$

To convey the intuition, we first study the equilibrium bidding strategies in the last auction $A_N^{(L)}$. Recall the following definitions:

- $M_N^{(L)} \equiv$ median of signals in a section of size $\Delta$ in $A_N^{(L)}$. 

37
• $\hat{M}_N^{(L)} \equiv$ median of bids in a section of size $\Delta$ in $A_N^{(L)}$, computed using the approximate market model.

• $\pi_l(x; F^{(N,L)}) \equiv$ equilibrium bidding strategy for a trader of type $l$, $l = 1, \ldots, k$ in $A_N^{(L)}$.

• $\pi_A(x; F^{(N,L)}) \equiv$ market equilibrium bidding strategy in $A_N^{(L)}$.

• $F_{\nu} A{\mathcal L}^{(N,L)}_i (x, m; F^{(N,L)}) \equiv \mathbb{E}_l \left[ V \mid X = x, M_N^{(L)} = m, F^{(N,L)} \right]$.

We will assume to begin with that each $\pi_l^{(i,j)}(x; F^{(i,j)})$ is strictly increasing in $x$ for all types in all auctions, an assertion which will be verified soon. Using this assumption we obtain:

**Lemma 2** The average (market) bidding strategy $\pi_A(x; F^{(N,L)})$ is strictly increasing in $x$, and hence $\hat{M}_N^{(L)} = \pi_A(M_N^{(L)}, F^{(N,L)})$.

**Proof.** Omitted.

Now consider an individual of type $l$ who receives a signal $x$ before the start of $A_N^{(L)}$. His problem is to:

$$\begin{align*}
\max_b & \quad \mathbb{E}_l \left[ (V - \hat{M})1_{\{b > \hat{M}\}} \mid X = x; F^{(N,L)} \right] \\
+ & \quad \mathbb{E}_l \left[ (\hat{M} - V)1_{\{b \leq \hat{M}\}} \mid X = x; F^{(N,L)} \right]
\end{align*}$$

or

$$\begin{align*}
\max_b & \quad \mathbb{E}_l \left[ (V - \hat{M})1_{\{b > \hat{M}\}} \mid X = x, M_N^{(L)} = m; F^{(N,L)} \right] \mid X = x; F^{(N,L)} \\
+ & \quad \mathbb{E}_l \left[ (\hat{M} - V)1_{\{b \leq \hat{M}\}} \mid X = x, M_N^{(L)} = m; F^{(N,L)} \right] \mid X = x; F^{(N,L)}
\end{align*}$$

Using lemma 2, this problem could be written as:

$$\begin{align*}
\max_b & \quad \mathbb{E}_l \left[ \mathbb{E}_l \left[ (V - \pi_A^{(N,L)}(m))1_{\{b > \hat{M}\}} \mid X = x, M_N^{(L)} = m; F^{(N,L)} \right] \mid X = x; F^{(N,L)} \right] \\
+ & \quad \mathbb{E}_l \left[ \mathbb{E}_l \left[ (\pi_A^{(N,L)}(m) - V))1_{\{b \leq \hat{M}\}} \mid X = x, M_N^{(L)} = m; F^{(N,L)} \right] \mid X = x; F^{(N,L)} \right]
\end{align*}$$

or

$$\begin{align*}
\max_b & \quad \mathbb{E}_l \left[ F_{\nu} A{\mathcal L}^{(N,L)}_i (x, m; F^{(N,L)}) - \pi_A^{(N,L)}(m) \right] \mid X = x; F^{(N,L)}
\end{align*}$$
\[ + \mathbb{E}_x \left[ \sum_{i \in I} \left( \pi_{A_i}^{(N,L)}(m) - \mathcal{F} \mathcal{V} \mathcal{A}_{I_i}^{(N,L)}(x, m; \mathcal{F}^{(N,L)}) \right) \right] \text{ or} \]

\[ \max_b \int_{-\infty}^{b_a^{-1}(b_1(z))} \left[ \mathcal{F} \mathcal{V} \mathcal{A}_{I_i}^{(N,L)}(x, m; \mathcal{F}^{(N,L)}) - \pi_{A_i}^{(N,L)}(m) \right] dF(m \mid x; \mathcal{F}^{(N,L)}) \]

\[ + \int_{b_a^{-1}(b_1(z))}^{\infty} \left[ \pi_{A_i}^{(N,L)}(m) - \mathcal{F} \mathcal{V} \mathcal{A}_{I_i}^{(N,L)}(x, m; \mathcal{F}^{(N,L)}) \right] dF(m \mid x; \mathcal{F}^{(N,L)}). \]

Where \( F(m \mid x; \mathcal{F}^{(N,L)}) \) is the conditional distribution of the median of signals in the last market section conditional on all information until the beginning of the last auction. Differentiating the objective function and imposing the condition that at equilibrium \( b = \pi_i(x; \mathcal{F}^{(N,L)}) \), we conclude that the first order necessary conditions for the solution of the above optimization problem is given by:

\[ \pi_i^{(N,L)}(x; \mathcal{F}^{(N,L)}) = \mathcal{F} \mathcal{V} \mathcal{A}_{I_i}^{(N,L)}(x, \pi_{A_i}^{1}(\pi_i(x; \mathcal{F}^{(N,L)})); \mathcal{F}^{(N,L)}) \quad \forall i = 1, \ldots, k \]

\[ \pi_{A_i}^{(N,L)}(x, \mathcal{F}^{(N,L)}) = \sum_{i=1}^{l=k} \pi_i^{(N,L)}(x, \mathcal{F}^{(N,L)}) \mu_{i}^{2} / \sum_{i=1}^{l=k} \mu_{i}^{2}. \] (9)

This is a system of \( k + 1 \) functional equations in \( k + 1 \) unknown functions \( \{\pi_1(x), \pi_2(x), \ldots, \pi_k(x), \pi_A(x)\} \). Existence of a solution of this system is recorded below.

**Proposition 7** There exists a solution to the system 9, in which \( \pi_i^{(N,L)}(x) \) and \( \pi_{A_i}^{(N,L)}(x) \) are increasing functions.

**Proof.** See Appendix B

Existence of a solution to the system (9) does not immediately imply that the anonymous double auction game has a solution. We need to check the participation constraint that the expected payoff to each type after observing \( x \) is nonnegative. We do this in the following proposition.

**Proposition 8 (Individual Rationality)** In the anonymous double auction game, the expected payoff to any type of players is nonnegative.
\textbf{Proof.} See Appendix B

Having established the existence of equilibrium in the last auction, we can now easily use the same arguments to establish the existence of equilibrium in the whole sequence of auctions. We prove this by backward induction. Consider the economy after arrival of information at time $t_{i+1}$, and assume that for any auction $A^{(j)}_{i+1}, j = 1, \ldots, L$, there exists an equilibrium in which the average (market) bidding strategy $\pi^{(i+1,j)}_A(x; \mathcal{F}^{(i+1,j)})$ is increasing in $x$.

Consider an individual of type $l$ who is to participate in auction $A^{(j)}_i$, an auction conducted after information renewal epoch $t_i$. From proposition 4, we know that this agent’s decision problem is to:

$$\max_{\pi^{(i,j)}_l} \mathbb{E}_l[(p_{i+1} - p^{(j)}_i) \mathcal{P} \mathcal{O} \mathcal{S}_i | X = x, \mathcal{F}^{(i,j)}].$$

But $p^{(j)}_{i+1} = \pi^{(i+1,j)}_A(M^{(j)}_{i+1})$, and hence:

$$\mathbb{E}_l[p_{i+1} | X = x, \mathcal{F}^{(i,j)}] = \mathbb{E}_l[\pi^{(i+1,j)}_A(M^{(j)}_{i+1}) | X = x, \mathcal{F}^{(i,j)}].$$

Define

$$\mathcal{FVA}_l^{(i,j)}(x, m; \mathcal{F}^{(i,j)}) \equiv \mathbb{E}_l[\pi^{(i+1,j)}_A(M^{(j)}_{i+1}) | X = x, M^{(j)}_i = m, \mathcal{F}^{(i,j)}].$$

From remark 4, $M^{(j)}_i$ is an increasing function of $V$, and thus it follows from proposition 5 that $\mathcal{FVA}_l^{(i,j)}(x, m; \mathcal{F}^{(i,j)})$ is increasing in both $x$ and $m$. Arguments similar to those used in the analysis of the last auction show that an individual’s problem is to:

$$\max_b \int_{-\infty}^{b^{-1}(b(z))} \left[ \mathcal{FVA}_l^{(i,j)}(x, m; \mathcal{F}^{(i,j)}) - \pi^{(i,j)}_A(m) \right] dF(m | x; \mathcal{F}^{(i,j)})$$

$$+ \int_{b(z)}^{\infty} \left[ \pi^{(i,j)}_A(m) - \mathcal{FVA}_l^{(i,j)}(x, m; \mathcal{F}^{(i,j)}) \right] dF(m | x; \mathcal{F}^{(i,j)}).$$

Where $F(m | x; \mathcal{F}^{(i,j)})$ is the distribution of the median of a market section taken after information arrival at $t_{i+1}$ conditional on the agent’s information received until auction $A^{(j)}_i$. The first order conditions for the agent’s problem are thus given by:
\[
\pi_l^{(i,j)}(x; \mathcal{F}^{(i,j)}) = \mathcal{F}\mathcal{V}\mathcal{A}\mathcal{L}_l^{(i,j)}(x, \pi_l^{-1}(\pi_l(x; \mathcal{F}^{(i,j)})); \mathcal{F}^{(i,j)}) \quad \forall l = 1, \ldots, k
\]

\[
\pi_l^{(i,j)}(x, \mathcal{F}^{(i,j)}) = \sum_{l=1}^{l=k} \pi_l^{(i,j)}(x, \mathcal{F}^{(i,j)}) \mu_l^2 / \sum_{l=1}^{l=k} \mu_l^2. \tag{10}
\]

Existence of a solution to this system of functional equations and the fact that the solution satisfies the individual rationality constraints can be established using arguments identical to those used in proving propositions 7 and 8. Since we have shown that equilibrium exists in the last auction \(A_{N}^{(L)}\), the induction argument used above shows existence of equilibrium in all auctions.

6 Analysis of Market Dynamics

In this section we discuss the relationship between trade volume and change in prices between two information update epochs \(t_i\) and \(t_{i+1}\). We first note that trade volume and market prices are random variables because of the random process by which traders in the economy get the chance to trade in the market. This randomness is additional to the basic uncertainty about future events. We, therefore, focus on the time series of expected volume and expected price change where the expectation is taken over all possible random selections of traders.

Second, we note that, aside from the random selection uncertainty, expectations of future endogenous parameters of the model will differ from one type of trader to another. After all, that is why they trade. We, hence, report results from the point of view of one fixed, though arbitrary, type of trader. The main conclusion of the section is that the expected volume of trade and the expected change in price between two information updates are related to two distinct properties of the pattern of information flow, and therefore there is no fundamental reason that there should be a particular relationship between these two variables.

The expected change in price is proportional to the expected change in the value of the median of signals between two periods. Expected trade volume, however, is related to the distribution of private information in the economy at \(t_i\) and \(t_{i+1}\). In our market, where the traded asset is of zero net supply, and where there is no new participants,
trade can only occur when an old seller decides to buy and simultaneously an old buyer decides to sell.

An old seller, who assumed his position earlier because he received information more pessimistic than average, will change his position only if he receives information more favorable than average, independent of the unconditional expected change in the value of the median of signals. This reversal of position is thus related to the reversals or fluctuations over time of a typical trader's private information around the average signal. Finally, we provide an example to show that the expected volume and expected change in price need not be fundamentally related.

**Proposition 9 (Expected Trade Volume)** For \( \alpha \in [0,1] \), let \( \mathcal{EPS}_i(\alpha) \) indicate the position of \( \alpha \)'s bid relative to the median of bids in the economy after arrival of information at time \( t_i \). Similarly, let \( \mathcal{SPS}_i(\alpha) \) indicate the position of \( \alpha \)'s signal relative to the median of signals in the economy at \( t_i \). In other words \( \mathcal{SPS}_i(\alpha) \) is +1 if \( \alpha \)'s signal is above the median and is −1 if it is below the median. Letting \( \text{VOL}_i \) denote the expected trade volume between \( t_i \) and \( t_{i+1} \), we then have:

\[
\text{VOL}_i = \lambda \left( \{ \alpha \in [0,1] : |\mathcal{EPS}_i(\alpha) - \mathcal{EPS}_{i-1}(\alpha)| \neq 0 \} \right)
\]

\[
= \lambda \left( \{ \alpha \in [0,1] : |\mathcal{SPS}_i(\alpha) - \mathcal{SPS}_{i-1}(\alpha)| \neq 0 \} \right).
\]

**Proof.** See Appendix B

Proposition 9 reveals the link between the nature of the distribution of information in the economy and trading volume in the period between two information updates. To characterize this relationship we introduce the following definition.
Definition 8 (Reversal Property of A Stochastic Process) Let \( \{X_i\}, i = 1, 2, \ldots \) be a discrete time stochastic process. The reversal of \( \{X_i\} \) at time \( i \), denoted \( R_i \), is defined as:

\[
R_i = \Pr(X_i \geq M_i \mid X_{i-1} \leq M_{i-1}),
\]

where \( M_i \) is the median of the random variable \( X_i \).

Remark 5 The reversal property of \( \{X_i\} \) at time \( t_i \) could also be defined as:

\[
R_i = \Pr(X_i \leq M_i \mid X_{i-1} \geq M_{i-1}).
\]

The reader can easily verify that this definition is equivalent to definition 8.

Remark 6 The reversal property of a process depends on the joint distribution of its component random variables. Therefore, two agents with two different information structures might have different evaluation of the reversal property of any process, and hence might have different expectations of the trading volume.

Proposition 10 For \( \alpha \in [0, 1] \), let \( \{X_i^\alpha\} \) be the process of private information flow, where we recall that \( \{X_i^\alpha\} \) are identically distributed for all \( \alpha \in [0, 1] \) and for all information structures. Let \( \{R_i\} \) be the sequence of reversals of \( \{X_i\} \), where we take \( R_1 \) to be equal to \( \frac{1}{2} \) by definition. For any information structure, we have the relation:

\[
VOL_i = R_i \quad i = 1, \ldots, N.
\]  

PROOF. Follows from definition 8 and proposition 9. \( \Box \)

Next, we consider expected change in prices. and from the point of view of one particular type of traders, define:

- \( \delta_{p_i} \equiv \mathbb{E}[p_{i+1}^{(j)} - p_i^{(j)}] = \mathbb{E}[\pi_A(m^{i+1}) - \pi_A(m^i)]; \)
- \( \sigma^2_{p_i} \equiv \text{var}[p_{i+1}^{(j)} - p_i^{(j)}] = \text{var}[\pi_A(m^{i+1}) - \pi_A(m^i)]; \)
- \( \delta_m = \mathbb{E}[m^{i+1} - m^i]; \)
- \( \sigma^2_m = \text{var}[m^{i+1} - m^i], \)
where \( m^i \) is the median of signals in a random section of size \( \Delta \) taken at time \( t_i \). The above expressions are the unconditional expected change in price and change in the median of signals; and the unconditional variance of these changes as seen by a trader of the chosen type. These quantities are computed without knowledge of the signals that a trader receives over time. We interpret these moments as those computed by a neutral observer who has no access to private signals and who follows the markets. He shares, however, the same “probabilistic views” as one type of the traders in the market. Such a neutral observer will come to the following conclusions:

**Proposition 11 (Price Change and Volatility)** For a particular type of neutral observers, and for all epochs \( t_i \), we have the following relations:

\[
\begin{align*}
\delta_{pij} &= a_{ij} \delta_m + b_{ij} \\
\sigma_{\pi ij}^2 &= c_{ij} \sigma_m^2 + d_{ij}
\end{align*}
\]

Where \( a_{ij}, b_{ij}, c_{ij}, d_{ij} \) are constants for all auctions \( A_t^{(j)} \), and where \( a_{ij} \) and \( c_{ij} \geq 0 \).

**Proof.** See Appendix B.

Our neutral observer concludes that the Market reacts in the manner given by proposition 11. The expected change in the median of signals leads to an expected change in prices with a market sensitivity factor \( a_{ij} \) for auction \( A_t^{(j)} \). Different types of observers think of the market sensitivity differently, and thus expect the market to react in different degrees for the same level of change in signals. One can also show, using the linkage principle of Milgrom and Weber (1982), that the sequence of numbers \( \{a_{i1}, a_{i2}, \ldots, a_{iL}\} \) will be increasing. Bidders in later auctions will respond more aggressively to changes in the level of signals because information revealed in earlier auctions weaken the so-called "winner's curse".

The above analysis shows that expected change in price and expected volume are related to two different properties of the information flow. We show in the following example that many possible relations between change in price and volume are possible. No particular relation is always valid for all futures contracts. Consider an environment in which at any time the private signal can take any one of four possible values. For simplicity of the exposition, let the possible values taken at time \( t_i \) be \( \{x_1, x_2, x_3, x_4\} \).
where $x_1 < x_2 < x_3 < x_4$. Let the possible values at $t_{i+1}$ be \{ $x'_1 = x_1$, $x'_2 = x_2 + s$, $x'_3 = x_3 + s$, $x'_4 = x_4$ \}, where $s$ can take either the value $\alpha$ or $\beta$ with equal probability. In addition $\alpha$ and $\beta$ are constrained such that $x'_1 < x'_2 < x'_3 < x'_4$. Let the corresponding transition probabilities be given by:

$$
T = \begin{bmatrix}
\frac{1}{2} - \frac{p}{2} & \frac{1}{2} - \frac{p}{2} & \frac{p}{2} & \frac{p}{2} \\
\frac{p}{2} & \frac{1}{2} - \frac{p}{2} & \frac{1}{2} - \frac{p}{2} & \frac{p}{2} \\
\frac{p}{2} & \frac{p}{2} & \frac{1}{2} - \frac{p}{2} & \frac{1}{2} - \frac{p}{2} \\
\frac{1}{2} - \frac{p}{2} & \frac{p}{2} & \frac{1}{2} - \frac{p}{2} & \frac{1}{2} - \frac{p}{2}
\end{bmatrix}
$$

Where $T_{jk} = \Pr(X_{i+1} = x'_k | X_i = x_j)$ for $j, k = 1, \ldots, 4$. We can easily see that:

- $VOL = p$
- $\delta_m = \frac{1}{2}(\alpha + \beta)$
- $\sigma_m^2 = \frac{1}{4}(\alpha - \beta)^2$

By independently varying the values of $p, \alpha, \beta$ over time, one can produce any pattern of relationship between the time series of changes in price and volume over the contract life $[0, T]$. For example, one could choose a sequence of values of $p, \alpha, \beta$ in which high values of $p$ occur simultaneously with negative values of $\alpha$ and $\beta$. thus producing a pattern in which expected price decrease is accompanied by high trading volume. On the other hand, one could choose an information structure in which the values of $p$ are stable over time, regardless of changes in $\alpha$ or $\beta$, thus producing a pattern of trade in which volume is fairly stable in spite of expected price changes. These results contrast with the existing theories of volume, for example Tauchen and Pitts (1983) and Karpoff (1985). In these theories, one assumes a relationship between the price and the trading patterns of agents with no “general equilibrium” support for such assumptions. In addition, the results are not linked to the nature of information flow in the Economy.

7 Concluding Remarks

This essay reports an exploratory and preliminary study which attempts to address the problem of price formation in a large market in the absence of liquidity traders. The
contributions in this essay come in three forms. First, we introduce a new formulation of the problem based on the concept of sequential anonymous games. We embed an explicit price formation mechanism, a double auction, in a general equilibrium model using a random mechanism that brings random samples of traders to the trading floor, where prices are determined. We believe that such an approach portrays some features of actual trading arrangements.

Second, we advance a notion of bounded rationality, in which traders use an approximate model of market response to form their bids. We justify this “approximation within a model” on the grounds of complexity and we defend it on the grounds of “robustness”.

Third, we prove existence of equilibrium in the market, and show that the equilibrium has a “no regret” property. In this equilibrium, traders condition their bids on their signals and on the fact that their bids will be executed. After learning the market price, no trader will regret his bid. We also characterize the relationship between the trading volume between two information updates in terms of the distribution of information in the economy. We show that expected price change and expected volume are related to two independent attributes of the information flow pattern, and there is no fundamental reason that a particular relation should consistently emerge between these two variables in all futures markets.

There is more work to be done. We need to incorporate the quantity a trader bids for as a strategic choice. In this essay, traders are constrained to take positions only in one contract. How the results change when traders can bid for a non-fixed number of contracts is an interesting open question. We also need to develop further the notion of bounded rationality. This essay shows that bounded rationality can be formally defined, justified and utilized. This is just one example, however. Developing other examples and hopefully some principles about a “reasonable” method of incorporating limits on computational and analytical abilities in a formal model should prove to be useful.

Finally, our work raises some questions about the empirical content of any result in an environment with heterogeneous agents. If agents interpret signals differently, they are bound to interpret endogenous variables also differently. A neutral economist, with no private information, who observes the market will not be able to construct or test models that all traders agree upon. He has to choose and fix a prior probability
measure for evaluation of the moments of the observed random variables before testing any hypothesis. How such tests are to be implemented, and how the results are to be interpreted is an important issue that warrants further research.

8 References


Appendix A  Asymmetric Double Auctions

The explicit price formation mechanism that we study within the framework of anonymous games is *Double Auctions*. A double auction is an exchange game with many buyers and many sellers of the same item. The equilibrium price results from the bids and offers made by buyers and sellers. This process contrasts with *Single Auctions* in which the price is determined by the competition on only one side of the market. This is the case of one seller against many buyers, or one buyer against many sellers.

The existing theory of auctions comprises two themes. First, the theory addresses the efficiency of auctions as economic institutions, and examines the allocations that result from auctions under the constraints of incentive compatibility. Technically, the analysis of such issues is carried out by invoking the *revelation principle*. As simplifying and powerful as the revelation principle is, its application leads to the disappearance of the specific strategic features of the exchange game. Those are replaced by revelation games in which strategies are merely reporting of private information and not taking market actions. Using this technique allows us to make certain statements about the efficiency of auctions. Yet, we do not get any answer concerning actual equilibrium strategies.

The second line of research in auction theory investigates different aspects of auctions by explicitly constructing the equilibrium strategies in the extensive form of the game induced by the auction structure. Having obtained such strategies, one could ask and provide answers to questions about the relationship between bidders’ private information/valuation and prices, about the sequence of prices in a sequence of auctions; and with the possibility of resale, about the relation between the price in the auction and that in the resale market.

In solving for equilibria in single auctions, models are, in most cases, formulated along lines of symmetry. In the case of a single seller, the buyers who have private information, are assumed to have valuations that depend in the same manner on exogenous factors, and each bidder’s valuation is assumed to be a symmetric function of other bidders’ signals. Moreover, the joint probability density of the agents’ signals is symmetric. In essence, each potential buyer is another copy of one representative trader. The only difference between traders is the realization of their private signals.

50
The symmetric formulation of auctions greatly simplifies the analysis of the bidding game. A bidder's strategy is a function relating how much he bids to the realization of his private signal. The search for equilibrium strategies is guided by the educated guess that there is a symmetric Nash equilibrium, in which all bidders adopt exactly the same strategy. This guess leads to the relatively easy identification of an equilibrium. Introducing asymmetry in a single auction game quickly complicates the analysis. The first order necessary conditions that the equilibrium strategies must satisfy are, in this case, a system of interrelated differential equations. Solving such a system of equations takes us afar into studying vector fields and away from obtaining clear economic intuition.

These difficulties are also encountered when one attempts to analyze a double auction even with symmetric buyers and sellers. To be sure, all the analyses in the literature study the equilibrium in symmetric double auctions in independent private valuation context. The value of the item to be exchanged is idiosyncratic to each agent, and there is no common value to the item.

One usually formulates a private valuation double auction as a Bayesian game with \( m \) buyers and \( n \) sellers. Buyers and sellers are risk neutral with inelastic supply (demand) for one unit of a certain item. All buyers are the same except for the realization of their independent private valuations, and similarly for the sellers. Buyers submit bids, which are the prices at or below which they are willing to buy one item, whereas sellers submit offers. A trader's strategy is a function relating his bid (offer) to his private valuation. Each auction has a pricing rule. A \( k \)-double auction is an auction in which the price is set to be equal to a convex combination of the lowest bid and highest offer that clear the market, with weights \( k \) and \( 1 - k \), where \( k \in [0, 1] \).

Technically, the analysis of a double auction with symmetric buyers and sellers results in a system of two interrelated differential equations that express the joint relationship between buyers and sellers' strategies, their derivatives, and the structural parameters of the model. As the analysis of Satterthwaite and Williams (1988) reveals, this system of differential equations has an infinite number of solutions. Moreover, the solutions are not obtained in a manner that conveys the intuition about the equilibria of the game.

Another attempt at solving this problem has been made by Williams (1988). In this approach the pricing rule in the market is set in such a way that one side of the market,
say the sellers, can not affect the price. The optimum strategy for a seller in this case is to bid his reservation price. Simplifying the problem in this manner, Williams (1988) obtained a characterization of the solution. His main focus was on the deviation of the buyers’ bids from their true valuation. An important insight obtained from this analysis is that as the number of traders increases, the difference between the traders bids (offers) and their private valuation diminishes, and thus the markets allow for capturing almost all gains from trade.

The current techniques and results in the auction literature are not suitable for the purpose of this work, namely, studying the dynamics of trading in a financial market using double auctions as the explicit price formation mechanism. The reason is two-fold. Double auctions have only been analyzed for the case of private valuations. Financial assets have no intrinsic consumption value, and their value is certainly the same for any individual. Thus, we need to develop a common value double auction model.

Second, in a market with buyers and sellers, the basic reason for trade is the differences in preferences, endowments, information and beliefs between traders. Therefore, a reasonable trading model in a financial market has to be an asymmetric model. The techniques and insights obtained in the analysis of common value single auctions have to be extended to arbitrarily asymmetric double auctions.
Appendix B  Proofs

PROOF OF PROPOSITION 3

PROOF. Let $\bar{z}$ be the median of $G$, and let $e \in E$, then $G(e)$ is given by:

$$G(e) = \sum_{i=1}^{i=k} \mu_i G_i(e)$$

$$= \sum_{i=1}^{i=k} \left[ G_i(e) - G_i(z_i^*) \right] + \sum_{i=1}^{i=k} \mu_i G_i(z_i^*)$$

$$= \sum_{i=1}^{i=k} \mu_i (e - z_i^*) g_i(\xi_i) + \frac{1}{2}$$

where $\xi_i \in [e, z_i^*)$ if $e \leq z_i^*$ and $\xi_i \in [z_i^*, e]$ if $e > z_i^*$. Let $\delta = G(e) - \frac{1}{2}$. It is clear that $\delta$ is a measure of the approximation error one introduces by using $e$ instead of the true median $\bar{z}$. Hence

$$\delta = \sum_{i=1}^{i=k} \mu_i (e - z_i^*) g_i(\xi_i).$$

For fixed $z_i^*, i = 1, \cdots, k$, $\delta$ is a linear operator from $\mathbb{R}^K$ to $\mathbb{R}$, with arguments $g_1(\xi_1), \cdots, g_k(\xi_k)$. For any estimator $e$, and for any constant $c > 0$, let $G_e(c)$ be the set of all $k$-tuples of distribution functions such that $\sum_i g_i^2(\xi_i) \leq c$. The norm $\|\delta_e\|$ of this operator gives the size of the maximum error possible over the set $G_e(1)$. In addition, the maximum possible error over the set $G_e(c) \leq c\|\delta_e\|$. Using the Euclidean metric, we get:

$$\|\delta_e\| = \sum_i \mu_i^2 (e - z_i^*)^2.$$ 

Let $e^*$, the robust estimator of $\bar{z}$ be given by the solution to:

$$\min_e \|\delta_e\|,$$

hence we conclude that $z^* \equiv e^* = \sum_{i=1}^{i=k} \mu_i^2 z_i^* / \sum_{i=1}^{i=k} \mu_i^2$. Note that for any $c > 0$ and for any other estimator $e$, we have that the maximum possible error on $G_{e^*}(c)$ is less than the maximum possible error on $G_e(c)$. Taking $c$ large enough, we conclude that $e^*$ minimizes the maximum possible error on the set of all $k$-tuple of distribution functions. $\blacksquare$

53
PROOF OF PROPPOSITION 4

PROOF. We first prove the necessity part. Write:

\[
\max_{\pi} E \left[ \sum_{i=1}^{i=N} (p_{i+1} - p_{i})\mathcal{P} \mathcal{O} \mathcal{S}_i \right] = \max_{\pi^{(i,j)}} \sum_{i=1}^{i=N} E \left[ E[(p_{i+1} - p_{i})\mathcal{P} \mathcal{O} \mathcal{S}_i \mid \mathcal{F}^{(i,j)}_a] \right]. \tag{15}
\]

Assume that \(\pi^*\) is a solution to the program given in (5), but that one of its components, say \(\pi^{(i,j)}\), is not a solution to (6). Then there exists another bidding strategy \(\hat{\pi}^{(i,j)}\) such that:

\[
E_{\hat{\pi}^{(i,j)}} \left[ (p_{i+1} - p_{i})\mathcal{P} \mathcal{O} \mathcal{S}_i \right] \geq E_{\pi^{(i,j)}} \left[ (p_{i+1} - p_{i})\mathcal{P} \mathcal{O} \mathcal{S}_i \right]. \tag{16}
\]

Substituting this into equation (15), we conclude that:

\[
E_{\hat{\pi}^*} \left[ \sum_{i=1}^{i=N} (p_{i+1} - p_{i})\mathcal{P} \mathcal{O} \mathcal{S}_i \right] \geq E_{\pi^*} \left[ \sum_{i=1}^{i=N} (p_{i+1} - p_{i})\mathcal{P} \mathcal{O} \mathcal{S}_i \right],
\]

where \(\hat{\pi}^*\) is the same bidding strategy as \(\pi^*\), except that at auction \(A_i^{(j)}\), \(\pi^{(i,j)}\) is replaced by \(\hat{\pi}^{(i,j)}\). Hence, we contradict the assumption that \(\pi^*\) is a solution to (5).

Now consider the sufficiency part. Recall equation (15). Since the value \(E[(p_{i+1} - p_{i})\mathcal{P} \mathcal{O} \mathcal{S}_i \mid \mathcal{F}^{(i,j)}_a]\) depends only on \(\mathcal{P} \mathcal{O} \mathcal{S}_i\), or equivalently on the bid \(\pi^{(i,j)}\), and since a trader’s bid in any auction has no effect on the market clearing prices in later auctions, then maximizing each element in the sum in the right-hand side in (15) is equivalent to maximizing the left-hand-side term.

PROOF OF LEMMA 1

PROOF. Computing \(\frac{d\rho(x)}{dx}\), we get:

\[
\frac{d\rho(x)}{dx} = \frac{F(x \mid v')f(x \mid v) - F(x \mid v)f(x \mid v') + [c_1f(x \mid v') - c_2f(x \mid v)]}{[F(x \mid v') - c_2]^2}
\]

But, by MLRP, we have:

\[
F(x \mid v') = \int_{-\infty}^{x} \frac{f(z \mid v')}{f(x \mid v)} f(z \mid v) \, dz \leq \frac{f(x \mid v')}{f(x \mid v)} F(x \mid v) \text{ or}
\]

54
\[ F(x \mid v')f(x \mid v) - F(x \mid v)f(x \mid v') \leq 0 \]  \hspace{1cm} (17)

Also using MLRP, we get:

\[ 1 - F(x \mid v') = \int_x^\infty \frac{f(z \mid v')}{f(z \mid v)} f(z \mid v) \, dz \geq \frac{f(x \mid v')}{f(x \mid v)} [1 - F(x \mid v)] \quad \text{or} \]

\[ F(x \mid v')f(x \mid v) - F(x \mid v)f(x \mid v') + [f(x \mid v') - f(x \mid v)] \leq 0 \]  \hspace{1cm} (18)

Using equations 17 and 18, and the fact that \(0 \leq c_1 \leq c_2 \leq 1\), we conclude that \(\frac{df}{dx} \leq 0\).
PROOF OF PROPOSITION 5

PROOF. 10 Recall that after the arrival of information at $t_i$, there is no aggregate uncertainty in the economy. In other words, we have, for all information structures:

$$\lambda\left(\{\alpha \in [0,1]: X_i^\alpha \leq z\} \mid v\right) = F(z \mid v).$$

Where we remind the reader that under assumption 5, $f(z \mid v)$ has the MLRP. The proof is in two steps. First, we will show that after taking random sections from $[0,1]$, the distribution of the remaining signals in the economy will still have the MLRP. Random slicing of the economy does not destroy this property. Second, we will show that an additional section from the economy will have a median whose distribution inherits the MLRP from the population from which it is taken.

To begin, we note the following simple relations. Let $A \subset [0,1]$ be a fixed set of measure $\lambda(A)$, and let $S$ be a random section of measure $\Delta$, then

$$\Pr(S \subset A) = \frac{[\lambda(A) - \Delta]^+}{1 - \Delta}$$

(19)

$$\Pr(S \subset A^c) = \frac{[1 - \lambda(A) - \Delta]^+}{1 - \Delta}$$

(20)

$$\Pr(\lambda(S \cap A) \leq \xi) = \frac{[1 - \lambda(A) - \Delta + \xi]^+}{1 - \Delta}$$

for $\xi \in [0, \lambda(A)]$  

(21)

For $s \in \mathbb{R}$ define

$$T_i^{(l)}(s) = \{\alpha \in [0,1]: X_i^\alpha \leq s \text{ and } \alpha \in \bigcup_{k=1}^{l-1} S_i^{(k)}\}$$

$$R_i^{(l)}(s) = \{\alpha \in [0,1]: X_i^\alpha \leq s \} \setminus T_i^{(l)}(s)$$

The set $T_i^{(l)}(s)$ is the set of all traders with signals not exceeding $s$ who have been already in the market before the start of auction $A_i^{(l)}$, whereas $R_i^{(l)}$ is the set of all traders with signals less than or equal to $s$, who are yet to participate in the market. Both $T_i^{(l)}$ and $R_i^{(l)}$ are random sets since their configurations depend on the configuration (or identity of agents) in the random sections $S_i^{(1)}, \ldots, S_i^{(l-1)}$.

10In this proof, and without loss of generality, we will drop the indices identifying information structures and those identifying the time epochs.
For any configuration of the sections \( S_i^{(1)}, \ldots, S_i^{(l-1)} \), and for all \( s \in \mathbb{R} \), let \( \Xi^{(l)}(s) \) be the random variable given by: \( \Xi^{(l)}(s) = \lambda(T_i^{(l)}(s) \setminus T_i^{(l-1)}(s)) \). Hence \( \lambda(R_i^{(l)}(s)) = F(s \mid v) - \sum_{k=1}^{l} \Xi^{(k)}(s) \). Define

\[
H^{(l)}(s \mid v) = \mathbb{E}[\lambda(R_i^{(l)}(s)) \mid v] = F(s \mid v) - \mathbb{E}[\sum_{k=1}^{l} \Xi^{(k)}(s) \mid v]
\] (22)

Where we note that the above expectation is taken over all possible configurations of the random sections \( S_i^{(1)}, \ldots, S_i^{(l-1)} \). We also note that \( \frac{H^{(l)}(s \mid v)}{1 - (l-1)\Delta} \) defines a probability distribution function on the real line, with the interpretation that \( \frac{H^{(l)}(s \mid v)}{1 - (l-1)\Delta} \) is the distribution function on the signals in the economy not yet brought to the market. We will use \( \frac{H^{(l)}(s \mid v)}{1 - (l-1)\Delta} \) to denote the Radon-Nikodym derivative of \( \frac{H^{(l)}(s \mid v)}{1 - (l-1)\Delta} \) with respect to the Lebesgue measure. Of course, \( H^{(1)}(s \mid v) \) is just \( F(s \mid v) \).

Our next step now is to show that for \( l = 1, \ldots, L \), the distribution functions \( \frac{H^{(l)}(s \mid v)}{1 - (l-1)\Delta} \) have the MLRP. We prove this by induction.

Assume that \( \frac{H^{(l)}(s \mid v)}{1 - (l-1)\Delta} \) has the monotone likelihood ratio property. Now take section \( S_i^{(l)} \), and assume that \( M_i^{(l)} = m_i \). Clearly:

\[
H^{(l+1)}(s \mid v) = H^{(l)}(s \mid v) - \mathbb{E}[\Xi^{(l+1)}(s \mid v)].
\] (23)

We proceed to compute \( H^{(l+1)}(s \mid v) \), for \( s \leq m_i \), which is given by:

\[
H^{(l+1)}(s \mid v) = H^{(l)}(s \mid v) \ast \Pr(\Xi^{(l+1)}(s \mid v) = 0) + \int_0^{\Delta/2} [H^{(l)}(s \mid v) - \xi^+] dG(\xi) + [H^{(l)}(s \mid v) - \Delta/2] \ast \Pr(\Xi^{(l+1)}(s \mid v) = \Delta/2),
\] (24)

where \( G(\xi \mid v) \) is the conditional distribution function on the measure of a random slice taken out of \( R_i^{(l)} \). Let \( c_v = [H^{(l)}(m_i \mid v) - \Delta/2] \). Using Equations 19, 20, and 21, we compute the following expressions for equation 24.

1. \( H^{(l)}(s \mid v) \leq \Delta/2 \) or \( s \in (-\infty, s_i^*(v)) \):

\[
c_v H^{(l+1)}(s \mid v) = H^{(l)}(s \mid v)[H^{(l)}(m_i \mid v) - H^{(l)}(s \mid v) - \Delta/2] + \int_0^{H^{(l)}(s \mid v)} [H^{(l)}(s \mid v) - \xi] d\xi
\] (25)

\[
c_v h^{(l+1)}(s \mid v) = h^{(l)}(s \mid v)[H^{(l)}(m_i \mid v) - H^{(l)}(s \mid v) - \Delta/2].
\] (26)
2. \( \Delta / 2 \leq H^{(l)}(s \mid v) \leq H^{(l)}(s \mid v) - \Delta / 2 \) or \( s \in (s_1^*(v), s_2^*(v)) \):

\[
c_v H^{l+1}(s \mid v) = H^{(l)}(s \mid v) [H^{(l)}(m_l \mid v) - H^{(l)}(s \mid v) - \Delta / 2] + \int_0^{\Delta / 2} [H^{(l)}(s \mid v) - \xi] d\xi + [H^{(l)}(s \mid v) - \Delta / 2]^2
\]

(27)

\[
c_v h^{l+1}(s \mid v) = h^{(l)}(s \mid v) [H^{(l)}(m_l \mid v) - \Delta].
\]

(28)

3. \( H^{(l)}(m_l \mid v) - \Delta / 2 \leq H^{(l)}(s \mid v) \leq H^{(l)}(m_l \mid v) \) or \( s \in (s_2^*(v), m_l) \):

\[
c_v H^{l+1}(s \mid v) = \int_{\Delta / 2 - [H^{(l)}(m_l \mid v) - H^{(l)}(s \mid v)]}^{\Delta / 2} [H^{(l)}(s \mid v) - \xi] d\xi + [H^{(l)}(s \mid v) - \Delta / 2]^2
\]

(29)

\[
c_v h^{l+1}(s \mid v) = h^{(l)}(s \mid v) [H^{(l)}(s \mid v) - \Delta / 2].
\]

(30)

Note that both \( H^{(l+1)}(s \mid v) \) and \( h^{(l+1)}(s \mid v) \) are both continuous on \( R \). Now we proceed to compute the likelihood ratio \( R = \frac{h^{(l+1)}(s \mid v)}{h^{(l+1)}(s \mid v')} \), for \( v' > v \). Note that by remark 4 we have that \( s_1^*(v) < s_1^*(v') \) for \( i = 1, 2 \). \( R \) is thus given by:

\[
R = \begin{cases} 
    c_1 \frac{H^{(l+1)}(s \mid v)}{H^{(l+1)}(s \mid v') - H^{(l+1)}(s \mid v) - \Delta / 2} & s \leq s_1^*(v) \\
    c_2 \frac{H^{(l+1)}(s \mid v)}{H^{(l+1)}(m_l \mid v) - H^{(l+1)}(s \mid v) - \Delta / 2} & s_1^*(v) \leq s \leq s_2^*(v') \\
    c_3 \frac{H^{(l+1)}(s \mid v)}{H^{(l+1)}(s \mid v')} & s_2^*(v') \leq s \leq s_2^*(v') \\
    c_4 \frac{H^{(l+1)}(s \mid v)}{H^{(l+1)}(s \mid v') - \Delta / 2} & s_2^*(v') \leq s \leq m_l
\end{cases}
\]

Where \( c_i, i = 1, \cdots, 5 \) are constants. On each of these intervals \( R \) is nonincreasing either by the assumption of MLRP or by virtue of lemma 1. This fact together with the continuity of the likelihood functions establish that \( R \) is nonincreasing on \((-\infty, m_l)\) for \( v' > v \). An identical proof shows that it is also nonincreasing on \([m_l, \infty)\). This shows that \( h^{(l+1)}(s \mid v) \) inherits the MLRP from \( h^{(l)}(s \mid v) \).

Now we prove the second step, namely that the median of a section of size \( \Delta \) taken from a population with a distribution of signals that has the MLRP will inherit this property. Let \( H^{(l)}(s \mid v) \) be the distribution of signals in the economy before auction \( A_i^{(l)} \). Take a section of size \( \Delta \). For \( s \) such that \( H^{(l)}(s \mid v) \geq \Delta / 2 \), we have

\[
Pr\left(M_i^{(l)} \leq s \mid v\right) \equiv G^{(l)}(s \mid v) = \frac{H^{(l)}(s \mid v) - \Delta / 2}{1 - 1/\Delta}
\]

and

58
\[ g^{(l)}(s \mid v) = \frac{h^{(l)}(s \mid v)}{1 - l\Delta}. \]

Hence the distribution of the median inherits the MLRP. Note that for the last auction, the probability density is concentrated at a single point, which is equal to the expectation of the median in the last segment conditional on the realization of medians in all previous segments. This completes our proof. \( \Box \)

**Proof of Proposition 7**

**Proof.** Consider the space \( \mathcal{H} \) of increasing functions from \( \mathbb{R} \to \mathbb{R}_+ \), which are bounded above by \( \nu_2 \) and bounded below by \( \nu_1 \). Let \( \mathcal{H}^{k+1} \) be the \( k + 1 \)-fold product of \( \mathcal{H} \). For \( h_1 \) and \( h_2 \in \mathcal{H}^{k+1} \), let:

\[
  h_1 \lor h_2(x) = \max[h_1(x), h_2(x)] \quad \text{and} \quad h_1 \land h_2(x) = \min[h_1(x), h_2(x)],
\]

where the minimum and maximum are taken component-wise. Define the order relation \( \succeq \) on \( \mathcal{H}^{k+1} \) as:

\[
  h_1 \succeq h_2 \quad \text{if and only if} \quad h_1^{(i)}(x) \geq h_2^{(i)}(x) \ \forall i = 1, \ldots, k + 1, \ \forall x \in \mathbb{R}.
\]

The space \( (\mathcal{H}^{k+1}, \succeq) \) is clearly a complete lattice. Now consider the mapping \( \varphi: \mathcal{H}^{k+1} \to \mathcal{H}^{k+1} \) given by the system of equations:

\[
  [\varphi(h)]^{(i)}(x) = \varphi_{\mathcal{F}, \mathcal{A}_c}(x, [h^{k+1}]^{-1}(h^{(i)}(x); \mathcal{F}^{(N,L)})) \quad i = 1, \ldots, k
\]

\[
  [\varphi(h)]^{(k+1)}(x) = \frac{\sum_{i=1}^{k} h^{(i)}(x) \mu_i^2}{\sum_{i=1}^{k} \mu_i^2}.
\]

From the properties of \( \varphi_{\mathcal{F}, \mathcal{A}_c}(\cdot, \cdot, \cdot) \) recorded in proposition 6, one concludes that \( \varphi \) is a monotonically increasing mapping from \( \mathcal{H}^{k+1} \) to itself. It then follows from Tarski's fixed point theorem (Tarski 1955) that there is a fixed point for \( \varphi \). A fixed point for \( \varphi \) is a solution to the system of functional equations given in (9). \( \Box \)

**Proof of Proposition 8**
PROOF. The expected utility for an individual of type \( l \), who has observed a signal \( x \) is given by:

\[
\int_{-\infty}^{\pi_A^{-1}(\pi_l(x))} \left[ \mathcal{F} \mathcal{V} \mathcal{A} \mathcal{L}_i^{(N,L)}(x, m; \mathcal{F}^{(N,L)}) - \pi_A(m) \right] \, dF(m \mid x; \mathcal{F}^{(N,L)}) + \\
\int_{\pi_A^{-1}(\pi_l(x))}^{\infty} \left[ \pi_A(m) - \mathcal{F} \mathcal{V} \mathcal{A} \mathcal{L}_i^{(N,L)}(x, m; \mathcal{F}^{(N,L)}) \right] \, dF(m \mid x; \mathcal{F}^{(N,L)}).
\]

From which we see that a sufficient condition for the nonnegativity of expected payoffs is:

\[
\mathcal{F} \mathcal{V} \mathcal{A} \mathcal{L}_i^{(N,L)}(x, m; \mathcal{F}^{(N,L)}) - \pi_A(x; \mathcal{F}^{(N,L)}) \geq (\geq) 0 \quad \text{as} \quad \pi_l(x; \mathcal{F}^{(N,L)}) \geq (\leq) \hat{m}.
\]

Recalling that \( \mathcal{V} \mathcal{A} \mathcal{L}_i(x, \hat{m}; \mathcal{F}^{(N,L)}) = \mathbb{E}_l\left[ V \mid X = x, \hat{M}_N^{(L)} = \hat{m}; \mathcal{F}^{(N,L)} \right] \), then we have:

\[
\mathcal{V} \mathcal{A} \mathcal{L}_i(x, \hat{m}; \mathcal{F}^{(N,L)}) = \mathcal{F} \mathcal{V} \mathcal{A} \mathcal{L}_i(x, \pi_A^{-1}(\hat{m}); \mathcal{F}^{(N,L)}),
\]

by virtue of the bounded rationality assumption. Now suppose that \( \pi_l(x) > \hat{m} \). By the monotonicity of \( \pi_l \), there exists a \( y < x \) such that \( \pi_l(y) = \hat{m} \). Now consider

\[
\mathcal{F} \mathcal{V} \mathcal{A} \mathcal{L}_i(x, m; \mathcal{F}^{(N,L)}) = \mathcal{V} \mathcal{A} \mathcal{L}_i(x, \hat{m}; \mathcal{F}^{(N,L)}) = \mathcal{F} \mathcal{V} \mathcal{A} \mathcal{L}_i(x, \pi_A^{-1}(\hat{m}); \mathcal{F}^{(N,L)}) = \\
g_i(x, \pi_A^{-1}(\pi_l(y)); \mathcal{F}^{(N,L)}) \geq g_i(y, \pi_A^{-1}(\pi_l(y)); \mathcal{F}^{(N,L)}) = \pi_l(y) = \hat{m}.
\]

Similarly for the case when \( \pi_l(x) \leq \hat{m} \). This proves our result. \( \blacksquare \)
Proof of Proposition 9

Proof. Fix a random selection of sections in the economy. Total volume between \( t_i \) and \( t_{i+1} \) is given by:

\[
V_i = \sum_{j=1}^{j=L} \lambda\left(\{\alpha \in S_i^j: |B\text{POS}_i(\alpha) - B\text{POS}_{i-1}(\alpha)| \neq 0\}\right)
\]

\[
= \lambda\left(\{\alpha \in [0, 1]: |B\text{POS}_i(\alpha) - B\text{POS}_{i-1}(\alpha)| \neq 0\}\right),
\]

since the market sections are disjoint. Now write:

\[B\text{POS}_i(\alpha) = A[\varepsilon\text{POS}_i(\alpha)],\]

where \( A \) is a random variable defined on the space of possible configurations of market sections and which can take the values \( \{-1, +1\} \). When \( A = 1 \), the bidder’s relative position in the market section is the same as his relative position in the economy. On the other hand, when \( A = -1 \), the random selection procedure results in a market relative position that does not reflect the trader’s position relative to the median of the economy. Now consider the set

\[B = \{\alpha \in [0, 1]: |B\text{POS}_i(\alpha) - B\text{POS}_{i-1}(\alpha)| \neq 0\}.
\]

\( B \) is a random set that depends on the configuration of sections. We would like to estimate the "average" measure of this set.

Let \( E \) denote the expectation operator over the probability space of all possible configurations. Fixing \( \alpha \in [0, 1] \), we note that:

\[E | B\text{POS}_i(\alpha) - B\text{POS}_{i-1}(\alpha)| = |\varepsilon\text{POS}_i(\alpha) - \varepsilon\text{POS}_{i-1}(\alpha)| \leq E(A).
\]

Under the assumption that \( E(A) \neq 0 \), we have:

\[E | B\text{POS}_i(\alpha) - B\text{POS}_{i-1}(\alpha)| = 0 \iff |\varepsilon\text{POS}_i(\alpha) - \varepsilon\text{POS}_{i-1}(\alpha)| = 0.
\]

Using Fubini’s theorem, we conclude that:

\[
VOL_i = E\left[\lambda\left(\{\alpha \in [0, 1]: |B\text{POS}_i(\alpha) - B\text{POS}_{i-1}(\alpha)| \neq 0\}\right)\right]
\]

\[
= \lambda\left(\{\alpha \in [0, 1]: E | B\text{POS}_i(\alpha) - B\text{POS}_{i-1}(\alpha)| \neq 0\}\right)
\]

\[
= \lambda\left(\{\alpha \in [0, 1]: |\varepsilon\text{POS}_i(\alpha) - \varepsilon\text{POS}_{i-1}(\alpha)| \neq 0\}\right).
\]
From the equilibrium strategies, we know that the bidding strategies are increasing in the value of the private signals, hence

$$\mathcal{EPS}_{i}(\alpha) = +1 \iff \pi_i(X^i) \geq \pi_A(M^i),$$

for all types $i$, and where $M^i$ is the median of signals in the economy at time $t_i$. This implies that

$$\sum_i \mu_i \pi_i(X^i) \geq \pi_A(M^i) \quad \text{or} \quad \pi_A(X^i) \geq \pi_A(M^i) \quad \text{or} \quad \mathcal{SPPOS}_i(\alpha) = \mathcal{EPS}_i(\alpha),$$

which completes our proof. \qed

**Proof of Proposition 11**

**Proof.**

First consider the case when the market bidding strategy $\pi_A^i(x)$ is differentiable. Write

$$\delta_{p_j} = E[\pi_A^i(m^{i+1}) - \pi_A^i(m^i)]$$

$$= E[\pi_A^i(\xi)(m^{i+1} - m^i)]$$

where $\pi_A^i(\xi)$ is the derivative of $\pi_A^i$, and where $\xi \in [m^i, m^{i+1}]$ if $m^{i+1} \geq m^i$, and $\xi \in [m^{i+1}, m^i]$ if $m^{i+1} \leq m^i$. Hence

$$\delta_{p_j} = a_j \delta_m + b_j \quad \text{where}$$

$$a_j = E[\pi_A^i(\xi)] \quad \text{and}$$

$$b_j = \text{cov}(m^{i+1} - m^i, \pi_A^i(\xi)).$$

Similarly,

$$\sigma_{p_j}^2 = c_j \sigma_m^2 + d_j \quad \text{where}$$

$$c_j = \text{var}[\pi_A^i(\xi)] \quad \text{and}$$

$$d_j = \text{cov}((m^{i+1} - m^i)^2, \pi_A^i(\xi)^2) - \left(\text{cov}(m^{i+1} - m^i, \pi_A^i(\xi))^2\right).$$
It is clear that both \( a_j \) and \( c_j \) are positive, since the market bidding strategy is increasing. When \( \pi_A^j(x) \) is not differentiable, simply substitute its derivative by the function:

\[
S(m^{i+1}, m^i) = \frac{\pi_A^j(m^{i+1}) - \pi_A^j(m^i)}{m^{i+1} - m^i}
\]

This function is positive, and the coefficient \( a_j \) is equal to its expectation over the values of the random variables \( m^i \) and \( m^{i+1} \).
Appendix C  Differential information in an economy with a continuum of agents

We here briefly show how to construct a model with a continuum of agents who observe different random variables. Feldman and Gilles(1985) have recently shown that it is not correct to postulate that the family of random variables observed by agents is a continuum of independent and identically distributed random variables.

In our model, the economy is the measure space \(([0, 1], \mathcal{B}[0, 1], \lambda)\). Let \((\Omega, \mathcal{F}, P)\) be a probability space on which are defined a family of i.i.d random variables, such that for all \(\alpha \in [0, 1]\), \(X_\alpha: \Omega \to \mathbb{R}\), with a given distribution function, say \(F(x)\). Now let the random function \(X: \Omega \times [0, 1] \to \mathbb{R}\) be defined by: \(X(\omega, \alpha) = X_\alpha(\omega)\). Feldman and Gilles (1985) proposition 1 shows that for every \(\omega \in \Omega\), either \(X(\omega, .)\) is not lebesgue measurable or there exists an open set \(G \in \mathcal{B}[0,1]\) such that the law of large number fails on \(G\). In other words:

\[
\int_G X(\omega, \alpha) \lambda(d\alpha) \neq E[X]\lambda(G).
\]

Feldman and Gilles(1985) propose some remedies for this situation. That is they propose alternative ways of constructing probabilistic models of distribution of random variables over a continuum of agents while maintaining the no-aggregates uncertainty property, a form of the law of large numbers. We adopt their solution in which the independence condition is relaxed. Our formulation is based on the result contained in proposition 2 in their article, which we reproduce here. The interested reader can consult Feldman and Gilles(1985) for further details.

**Proposition 12 (Feldman and Gilles(1985))** Let \((S, \mathcal{L}, \nu)\) be a probability space with \(S\) a complete seperable metric space and \(\mathcal{L}\) the Borel \(\sigma\)-field. Let \(A = [0, 1]\) be the index set of individuals, \(\mathcal{B}\) be the Borel \(\sigma\)-field on \(A\) and \(\lambda\) be the Lebesgue measure. There exists a probability space \((\Omega, \mathcal{F}, P)\) which may be taken to be \(([0, 1], \mathcal{B}[0,1], \lambda)\), and a family of \(S\)-valued random variables \(\{X_\alpha: \alpha \in A\}\) such that:

1. For all \(\alpha \in A\), \(X_\alpha\) has distribution \(\nu\), i.e. \(P\{X_\alpha \in D\} = \nu(D)\) for all \(D \in \mathcal{L}\).

2. For all \(\omega \in \Omega\), \(\lambda(\{\alpha \in A: X_\alpha(\omega) \in D\}) = \nu(D)\) for all \(D \in \mathcal{L}\).
Essay II

OPTIMAL CONSUMPTION WITH INTERTEMPORAL SUBSTITUTION I:
THE CASE OF CERTAINTY

Abstract

We study the problem of optimal consumption choice in continuous time under certainty for a class of utility functions that capture the notion that consumptions at nearby dates are almost perfect substitutes. The class we consider excludes all time-additive and almost all the non time-additive utility functions used in the literature. We provide necessary and sufficient conditions for a consumption policy to be optimal. Furthermore, we demonstrate our general theory by solving in a closed form the optimal consumption policy for a particular felicity function. The optimal policy in our solution consists of a (possible) initial "gulp" of consumption, or an initial period of no consumption, followed by consumption at the rate that maintains a constant ratio of wealth to average past consumption.
1 Introduction

The choice of an optimal consumption plan for an individual under certainty is a classical problem in economics; see, for example, Ramsey (1928) and Modigliani and Brumberg (1954). In a large part of the literature that addresses this problem, the preferences of an individual are represented by a time-additive utility function. It has long been recognized that a time-additive utility function does not incorporate the intuitively appealing notion that past consumption can contribute to an individual's future satisfaction. Nevertheless, this class of preferences has been utilized because of its tractability. In discrete time models, one can argue that if the length of time periods is chosen appropriately, then past consumption would have negligible effect on current satisfaction. Such arguments, however, no longer apply when we consider a continuous time model.

Recently, a class of non-time-additive utility functions, due originally to Koopmans (1960) and Uzawa (1968), has gained increasing significance in the literature of continuous time equilibrium models; see, for example, Epstein (1987). Also, non-time-additive utility functions that exhibit habit formation are objects of current research interest in continuous time financial asset pricing models; see, for example, Constantinides (1988), Heaton (1988), and Sundaresan (1989). A common feature of these utility functions is that the felicity, or the index of instantaneous satisfaction at any time, is a function of the current consumption and a certain functional of the past consumption. A high level of past consumption can serve either to depress current appetite as in the case of the Uzawa type utility functions or to increase current appetite as in the case of the habit formation models.

These non-time-additive utility functions offer intuitively appealing notions of interactions between past and current consumption. They are also mathematically tractable. Optimal consumption policies can be characterized using dynamic programming. For particular functional forms of the felicity function, one can even have closed form expressions for these policies. Unfortunately, these utility functions fail to exhibit one key economic property that underlies much of our intuition about the behavior of asset prices over time in the absence of arbitrage possibilities.

A common hypothesis maintained by many financial economists is that asset prices
cannot have predictable jumps. To support this hypothesis, the argument goes as follows: suppose agents knew that the price of an asset were to increase by a discrete amount at a specific moment. If an agent could borrow funds for a very short period of time over which the interest he would have to pay is negligible, then he would be able to realize an "arbitrage profit". Having access to a frictionless market, he would simply borrow some funds just before the jump in the price were supposed to occur, buy the asset, wait for its price to increase, and then liquidate his position. The demand of agents exploiting this arbitrage opportunity would push the price of the asset up before the time when the jump were supposed to appear and hence the jump would be eliminated. The validity of this argument depends crucially on the hidden assumption that prices for consumption at nearly adjacent dates are almost equal and thus the interest that the agent pays on funds borrowed over a short time interval is negligible.

There are several economic forces that will bring about the closeness of prices for consumptions at nearly adjacent dates. The most primitive among these forces is the hypothesis that agents treat consumptions at nearly adjacent dates as almost perfect substitutes and thus ensure that almost equal prices are established for them. Huang and Kreps (1989), in a model under certainty, and Hindy and Huang (1989), in a model under uncertainty, study preferences that exhibit this property. They show in particular that if the felicity function depends explicitly on current consumption and is strictly concave in it, then consumptions at nearly adjacent dates cannot be treated as almost perfect substitutes. Since the functional forms of the felicity function used both in the Uzawa type and in the habit formation type preferences are strictly concave, agents with those preferences do not view consumptions at nearby dates as close substitutes. Therefore, such models will reach economic conclusions that violate some of our basic intuition about price behavior over time.

The purpose of this essay is to investigate the optimal consumption problem under certainty for a class of utility functions that treat consumptions at nearby dates as close substitutes. The key feature of this class of utility functions is that the felicity function at any time depends only upon an exponentially weighted average of past consumption – one derives satisfaction only from past consumption. We will demonstrate the tractability of this family by showing how dynamic programming can be used to give sufficient con-
ditions for optimality. In particular, for a specific functional form of the felicity function, we solve in closed form the optimal consumption policy.

Intuition suggests that when one derives current satisfaction only from past consumption, the optimal consumption policy may specify periods of consumption mixed with periods of no consumption. Our closed form solution exhibits this property. In particular, the optimal consumption policy in our solution calls for a (possible) initial "gulp" of consumption, or an initial period of no consumption, followed by consumption at the rate that maintains a constant ratio of wealth to average past consumption. The consumption rate is determined by the interplay between the effect of past consumption, the current return on investment and the impatience of the agent.

The rest of this essay is organized as follows. Section 2 sets up the consumption problem under certainty in continuous time. Section 3 heuristically derives necessary conditions for a consumption policy to be optimal. Section 4 provides sufficient conditions for optimality. In section 5, we solve the optimal consumption problem in closed form for a particular felicity function. Section 6 contains concluding remarks.

2 The Setup

Consider an economic agent who lives from time $t = 0$ to $t = T$ in a certain world where there is a single consumption good available at any time between 0 and $T$. The agent can consume "gulps" of the good at any moment, and can consume at finite rates over intervals. He can also refrain from consumption for any period as he sees fit. We represent the agent's consumption pattern over his life span by a positive\(^1\), increasing, right continuous function $C: [0, T] \to \mathbb{R}_+$, with $C(t)$ denoting the accumulated consumption from time zero to time $t$. Note that the only possible kind of a discontinuity of $C$ is a jump. Note also that we allow $C$ to have a "singular" component, that is a continuous increasing function whose derivative is zero for almost all $t$. The famous Cantor function is an example; see, for example, Royden (1968, p.48). The consumption set of the agent, $X_+$, is the space of all such functions and the commodity space, $X$, is the linear span of $X_+$. Note that $X$ is the space of right-continuous and finite variation functions. Note

\(^1\)We use weak relations in our discussion. Hence increasing is equivalent to nondecreasing, for example. When a relation is strict, we will explicitly state so.
also that a finite variation function $x$ on $[0,T]$ has a finite left-limit at any $t \in (0,T]$ denoted by $x(t^-)$ and a finite right-limit at any $t \in [0,T)$ denoted by $x(t^+)$. By right-continuity, for any $x \in X$ we have $x(t^+) = x(t)$, $t \in [0,T)$. For convenience, we will use the convention that $x(0^-) = 0$. Since left-limits exist for any $x \in X$, a jump of of $x$ at $t$ is $\Delta x(t) \equiv x(t) - x(t^-)$.

The points of discontinuity of a consumption pattern $C$ are the moments when the agent decides to consume a “gulp”, and the size of the jump is the amount of that “gulp”. Note that there can be only at most a countable number of times of “gulps” since $C$ is an increasing right continuous function of time. Let these moments be $\tau_1, \tau_2, \ldots$, and let the jump at time $\tau_i$ be $\Delta C(\tau_i)$. Moreover, $C$ is absolutely continuous on the intervals on which the agent consumes at rates and the derivative of $C$ on these intervals, denoted $c(t)$, is the “rate” of consumption there. When this derivative is zero, $C$ is constant on the interval, and the agent is not consuming at all.

The agent starts at time 0 with endowment $W(0)$, and at any time $t$ he can invest part of or all his wealth in a riskless asset with instantaneous rate of return $r(t)$. We assume that the interest rate $r(t)$ is a continuous function of time. The agent’s wealth before consumption at time $t$, denoted $W(t)$, changes according to the following dynamics:

$$
\begin{align*}
\frac{dW(t)}{dt} &= W(t) r(t) dt - dC(t) \quad \text{for } t \in (\tau_i, \tau_{i+1}) \quad \text{for all } i, \\
W(\tau_i^+) &= W(\tau_i) - \Delta C(\tau_i) \quad \text{for all } \tau_i.
\end{align*}
$$

(1)

Note that $W(t)$ is a left continuous function and (1) is just the budget constraint over time. We impose the condition that $W(t) \geq 0$ for all $t \in [0,T]$.

The agent has preferences for consumption over the period $[0,T]$ and for final wealth $W(T)$.

\footnote{The astute reader might have observed that the final wealth should be $W(T^+)$, since wealth is a left continuous function of time. To simplify notation, we will use the slightly inaccurate $W(T)$ for final wealth. Our justification is that consumption gulps at $T$ add nothing to the satisfaction of the agent. Hence, at the optimal solution $W(T) = W(T^+)$.} He feels, however, that consumption at one time is a substitute for consumption at other, near-by times. Therefore, delaying or advancing consumption for a very small period of time has a very small effect on his total satisfaction. Huang and Kreps (1989), for the case of certainty, and Hindy and Huang (1989), for the case of uncertainty, formalize the idea that consumption preferences can be similar to those expressed by the agent. In particular, the authors advance a family of topologies on the space of
consumption patterns over time that capture the notion that consumptions at nearly adjacent dates are almost perfect substitutes. In this essay, we use one member of this family given as follows. Fix $p \geq 1$. Define the topology $T$ on the commodity space $X$ to be the topology generated by the norm: \footnote{The special treatment of $C(T)$ in the norm accounts for the fact that two consumption patterns which agree on $[0,T]$ but disagree on $T$ could not be distinguished if the norm were without the term involving $C(T)$.}

$$
\|C\| = \left( \int_0^T |C(t)|^p \, dt + |C(T)|^p \right)^{1/p}.
$$

Moreover, Huang and Kreps (1989) and Hindy and Huang (1989) show that there are intuitively appealing functional forms of utility which express preferences for consumption with intertemporal substitution properties. The agent has one of them. Fix any consumption pattern $C$ and a corresponding terminal wealth $W(T)$. At any time $t$, define the exponentially weighted average of past consumption, $z(t)$, by:

$$
z(t) = z(0^-)e^{-\beta t} + \beta \int_{0^-}^t e^{-\beta(t-s)} \, dC(s),
$$

where $z(0^-) \geq 0$ is a constant and $\beta$ is a weighting factor. Note that the lower limit of the integral in (3) is 0$, to account for the possible jump of $C$ at $t = 0$, and that $z(t)$ is a right continuous function. Moreover, $z$ has a jump exactly when $C$ does and $z$ has a singular component when $C$ does. Also observe that higher values of $\beta$ imply higher emphasis on the recent past and less emphasis on consumption in the distant past. Note also that the average past consumption $z(t)$ changes according to the following dynamics:

$$
dz(t) = \beta [dC(t) - z(t) \, dt] \quad \text{for } t \in (\tau_i, \tau_{i+1}) \quad \text{for all } i,
$$

$$
z(\tau_i) = z(\tau_i^-) + \beta \Delta C(\tau_i) \quad \text{for all } \tau_i.
$$

Given $z(0^-)$, the agent's utility for the consumption pattern $C$ and final wealth $W(T)$ is given by:

$$
U(z(0^-), 0; (C, W(T))) = \int_0^T u(z(t), t) \, dt + V(W(T)),
$$

where both $u: \mathbb{R}_+ \times [0,T] \to \mathbb{R} \cup \{-\infty\}$ and $V: \mathbb{R}_+ \to \mathbb{R} \cup \{-\infty\}$ are continuous and increasing functions. Furthermore, $V$ is strictly concave and $u$ is strictly concave in its first argument. Note that we have allowed the possibility that $u$ or $V$ can take the value $-\infty$ at zero. For example, $V(W) = \frac{1}{\alpha} W^{\alpha}$, where $\alpha < 0$. The function $u$ is, in
the terminology of Arrow and Kurz (1970), the felicity function that assigns the level of satisfaction derived from past consumption and $V$ is the bequeath function. Preferences given by (3) and (5) are continuous in the product topology generated by $T$ and the Euclidean topology on $\mathbb{R}_+$; see Hindy and Huang (1989), proposition 17.

Starting from a given level of wealth, each consumption plan $C$ determines, via the budget constraint, the value of terminal wealth $W(T)$. It also determines for every time $t$, the wealth $W(t)$, and the average past consumption $z(t^-)$. Note that at any $t$, the value $z(t^-)$ is the average past consumption up till $t$, excluding the possible consumption at $t$. We will call $W(t)$ and $z(t^-)$ the state variables at $t$, since they represent the status of the agent before making his consumption decision at $t$. The agent faces the problem of choosing a consumption pattern $C^* \text{ -- the optimal control -- from the space } X_+ \text{ -- the admissible controls -- }$ to maximize his utility, $U(z(0^-), 0; (C, W(T)))$, subject to the dynamics given in (1) and (4), and given that his initial wealth is $W(0)$ and given $z(0^-)$.

Formally, let $\mathcal{A}(W(0), 0) \subset X_+ \times \mathbb{R}_+$ be the space of all pairs of consumption pattern $C$ and final wealth $W(T)$ that satisfy the budget constraint of (1) starting with $W(0)$. The agent’s problem is to find:

$$
\sup_{(C, W(T)) \in \mathcal{A}(W(0), 0)} U(z(0^-), 0; (C, W(T))).
$$

(6)

A solution to (6) exists if the supremum is finite and is attained by some $(C^*, W^*(T)) \in \mathcal{A}(W(0), 0)$.

At times, we will also consider a sub-problem of (6) at some $t \in [0, T)$: given $\{C(s); s \in [0, t^-]\}$ and $W(t)$ determined by (1) (the budget constraint), solve

$$
\sup_{(C^t, W(T)) \in \mathcal{A}(W(t), t)} U(z(t^-), t; C^t, W(T)) \equiv \int_t^T u(z(s), s) ds + V(W(T)),
$$

(7)

where $C^t$ is an increasing and right-continuous function on $[t, T]$ representing accumulated consumption starting from $t$ with the convention that $C^t(s) = 0$ for all $s < t$, where $\mathcal{A}(W(t), t)$ denotes the consumption pattern $C^t$ and final wealth $W(T)$ that satisfy the budget constraint (1) on $[t, T]$ with an initial value $W(t)$, where

$$
z(s) = z(t^-)e^{-\beta(t-s)} + \int_{t^-}^s e^{-\beta(s-\xi)} dC^t(\xi) \text{ for } s \geq t.
$$

(8)
It is clear that once $W(t)$ and $z(t^-)$ are known, \{$C(s); s \in [0, t^-]$\} has no impact on the choices of $C^t$ in (7) and this justifies the notation $U(z(t^-), t; C^t, W(t))$. Moreover, if $(C^*, W^*(T))$ is a solution to (6), then \{$C^t(s) = C^*(s) - C^*(t^-); s \in [t, T]$\} is a solution to (7) with $z(t^-)$ and $W(t)$ corresponding to $(C^*, W^*(T))$.

3 Heuristic Derivation of Necessary Conditions

We now use heuristic arguments to derive a set of necessary conditions for a solution to (6). These conditions will imply that an optimal consumption pattern can have gulps only at $t = 0$.

Suppose $(C^*, W^*(T))$ is an optimal solution to (6) and let the corresponding wealth and average past consumption, or the state variables, be $(W^*(t), z^*(t^-))$ for $t \in [0, T]$. At any time $t$ and starting with wealth $W(t)$ and average past consumption $z(t^-)$, let the value function $J(W(t), z(t^-), t)$ be the maximum possible attainable satisfaction over the remaining period $[t, T]$. In other words, let:

$$J(W(t), z(t^-), t) = \sup_{(C^t, W(T)) \in \mathcal{A}(W(t), t)} \left[ \int_t^T u(z(s), s) \, ds + V(W(T)) \right],$$

where \{$z(s); s \in [t, T]$\} is defined according to (8).

Assume that $J$ is finite and differentiable in $W$ and $z$. The marginal value of wealth at time $t$, $J_W(W, z, t) \equiv \frac{\partial J(W, z, t)}{\partial W}$, measures the changes in the maximum attainable satisfaction that results from a very small change in wealth when everything else is kept constant. Similarly, the marginal value of the average past consumption $z$ at time $t$, $J_z(W, z, t) \equiv \frac{\partial J(W, z, t)}{\partial z}$, measures the change in the value function for infinitesimal changes in $z$, ceteris paribus.

Along the optimal solution, $J$ should satisfy the following conditions:

1 – Bellman Optimality Principle
The Bellman optimality principle; see, for example, Fleming and Rishel (1975), implies that for all times $t$ and $\tau \in [0, T]$ such that $t < \tau$,

$$J(W^*(t), z^*(t^-), t) = \int_t^\tau u(z^*(s), s) \, ds + J(W^*(\tau), z^*(\tau^-), \tau).$$

72
That is: an optimal policy on \([t, T]\) remains to be optimal over any subperiod \([\tau, T]\). In particular, this principle implies that along the optimal solution, the value function \(J(W^*, z^*, t)\) is a continuous function of time, even at times of jumps, or

\[ J(W^*(\tau), z^*(\tau^-), \tau) = \lim_{\Delta \to 0} J(W^*(\tau+\Delta), z^*(\tau+\Delta), \tau+\Delta) \equiv J(W^*(\tau^+), z^*(\tau), \tau) \quad \text{for all } \tau. \tag{11} \]

For brevity of notation, we will use \(\bar{J}(t)\) to denote \(J(W^*(t), z^*(t^-), t)\) and similarly use \(\bar{J}_W(t)\) and \(\bar{J}_z(t)\) to denote the partial derivatives of \(\bar{J}(t)\) with respect to \(W\) and \(z\), respectively.

2 - Continuous Marginal Value of \(W\) and \(z\)

Along the optimal policy, given the interest rates, the marginal value of \(\epsilon\) units of wealth at time \(t \in [0, T]\) must be equal to the marginal value of \(\epsilon e^{\int_t^{t+\Delta t} r(s)ds}\) units of wealth at time \(t + \Delta t\) for small \(\Delta t > 0\) or else some changes in \(C^*\) will be called for to change \(W^*\) and \(z^*\). That is, we must have

\[ \bar{J}_W(t)\epsilon = \bar{J}_W(t + \Delta t)\epsilon e^{\int_t^{t+\Delta t} r(s)ds}. \tag{12} \]

Letting \(\Delta t \downarrow 0\), we get \(\bar{J}_W(t) = \bar{J}_W(t^+)\). Thus \(\bar{J}_W(t)\) must be right-continuous on \([0, T]\). Along the same line of arguments, we show that \(\bar{J}_W(t)\) must be left-continuous on \((0, T]\) and thus \(\bar{J}_W(t)\) is continuous on \([0, T]\).

Similarly, by the Bellman optimality principle, a marginal value of \(\epsilon\) increase in \(z(t^-)\) at \(t\) must be equal to its marginal contribution to the felicity function on \([t, t + \Delta t]\) and to the value function at time \(t + \Delta t\). That is,

\[ \bar{J}_z(t)\epsilon = \int_t^{t+\Delta t} u_z(z^*(s), s)\epsilon e^{-\beta(s-t)}ds + \bar{J}_z(t + \Delta t)\epsilon e^{-\beta\Delta t}. \tag{13} \]

Letting \(\Delta t \to 0\), we conclude that \(\bar{J}_z(t)\) must be right-continuous on \([0, T]\). The left-continuity of \(\bar{J}_z(t)\) is established by noting that

\[ \bar{J}_z(t - \Delta t)\epsilon = \int_t^{-\Delta t} u_z(z^*(s), s)\epsilon e^{-\beta(s-t-\Delta t)}ds + \bar{J}_z(t)\epsilon e^{-\beta\Delta t} \tag{14} \]

and letting \(\Delta t \to 0\).

\(^4\text{We use } u_z(z) \text{ to denote } \frac{\partial u(z)}{\partial z}.\)
3- Dynamics of Marginal Values
By the hypothesis that \( r(t) \) is continuous in \( t \), (12) implies, as \( \Delta t \to 0 \), that

\[
\frac{d\bar{J}_W(t)}{dt} = -r(t)\bar{J}_W(t).
\]

Next note that if \( z^*(t) \) is continuous on some time interval \((t - \Delta t, t + \Delta t)\), then

\[
\int_{t+\Delta t}^{t} u_z(z^*(s), s)e^{-\beta(s-t)}ds \to u_z(z^*(t), t)
\]
and

\[
\int_{t-\Delta t}^{t} u_z(z^*(s), s)e^{-\beta(s-t+\Delta t)}ds \to u_z(z^*(t), t).
\]
Hence, in a time interval over which \( C^* \) has no "gulps", (13) implies, as \( \Delta t \to 0 \), that

\[
\frac{d\bar{J}_z(t)}{dt} = -u_z(z^*(t), t) + \beta \bar{J}_z(t).
\]

Also note that at \( t = T \), the marginal value of wealth, \( \bar{J}_W(T) \) is exactly equal to the derivative of \( V \) evaluated at the terminal wealth \( W^*(T) \). In addition, any unit of the good consumed at \( t = T \) has no effect on the agent's total satisfaction and hence \( \bar{J}_z(T) = 0 \).

4- Appropriate Times for "Gulps"
Applying Bellman's optimality principle in (10) on an interval \((\tau_i, \tau_{i+1})\) when our solution prescribes consumption at rates, we get the following Bellman equation:

\[
\max_{c(t)} \{ u(z^*(t), t) + \bar{J}_W[r(t)W^*(t) - c(t)] + \bar{J}_z\beta[c(t) - z^*(t)] + \bar{J}_t \} = 0 \quad \text{for all } t \in (\tau_i, \tau_{i+1}).
\]

(15)

Note that the Bellman equation is linear in the consumption rate \( c(t) \). Hence, the consumption rate that satisfies (15) is given by:

\[
\bar{c}(t) = 0 \quad \text{when} \quad \bar{J}_W(t) - \beta \bar{J}_z(t) > 0,
\]

\[
\bar{c}(t) \in [0, \infty) \quad \text{when} \quad \bar{J}_W(t) - \beta \bar{J}_z(t) = 0,
\]

\[
\bar{c}(t) = \infty \quad \text{when} \quad \bar{J}_W(t) - \beta \bar{J}_z(t) < 0.
\]

Bellman equation, therefore, suggests that over the periods in which consumption is at rates, it cannot be the case that (\( \beta \) times) the marginal value of \( z \) is greater than the
marginal value of \( W \). It also suggests that a “gulp” of consumption might be prescribed at any moment \( \tau \) when \( \bar{J}_W(\tau) - \beta \bar{J}_z(\tau) \leq 0 \). Suppose that we prescribe a “gulp” at \( \tau \) when this quantity is strictly negative. Since the marginal values of \( W \) and \( z \) are continuous, the function \( \bar{J}_W(t) - \beta \bar{J}_z(t) \) will be strictly negative on an interval \( (\tau - \epsilon, \tau + \epsilon) \) for some \( \epsilon > 0 \). Hence our candidate policy should prescribe a jump for all points \( t \in (\tau - \epsilon, \tau + \epsilon) \). Since this cannot happen as there can only be a countable number of jumps, our policy might prescribe a “gulp” only at exactly the first moment \( \tau \), on this particular interval, such that \( \bar{J}_W(\tau) - \beta \bar{J}_z(\tau) = 0 \).

5- No “Gulps” after \( t = 0 \)

Now suppose that our policy calls for a “gulp” of size \( \Delta \) at time \( \tau > 0 \). Note that by right continuity of \( C \), we cannot have two successive jumps at any one moment. This, together with the continuity of the marginal values of \( W \) and \( z \), implies that there are two strictly positive numbers \( \epsilon_1, \epsilon_2 \) such that:

\[
\bar{J}_W(t) - \beta \bar{J}_z(t) \geq 0 \quad \text{on} \quad (\tau - \epsilon_1, \tau),
\]

\[
\bar{J}_W(t) - \beta \bar{J}_z(t) = 0 \quad \text{when} \quad t = \tau,
\]

\[
\bar{J}_W(t) - \beta \bar{J}_z(t) \geq 0 \quad \text{on} \quad (\tau, \tau + \epsilon_2).
\]

Assuming that \( \bar{J}_W \) and \( \bar{J}_z \) are differentiable functions of time, this implies that \( \bar{J}_W(t) - \beta \bar{J}_z(t) \) is decreasing just to the left of \( \tau \) and increasing just to the right of \( \tau \). This condition, however, cannot hold for any size of jump \( \Delta \). From the dynamics of \( \bar{J}_W \) and \( \bar{J}_z \), we get that:

\[
\frac{d}{dt}[\bar{J}_W(\tau^+) - \beta \bar{J}_z(\tau^+)] - \frac{d}{dt}[\bar{J}_W(\tau) - \beta \bar{J}_z(\tau)] = \beta \left[ u_s(z^*(\tau) + \Delta, \tau) - u_s(z^*(\tau), \tau) \right].
\] (16)

But by strict concavity of the felicity function in \( z \), this last quantity is strictly negative. Hence, if \( \frac{d}{dt}[\bar{J}_W - \beta \bar{J}_z] \) were negative just before the jump, it would be strictly negative just after the jump, thus violating the condition that \( \bar{J}_W - \beta \bar{J}_z \) be increasing just after a “gulp”. Note that our analysis is not valid for \( t = 0 \) and \( t = T \), thus \( C^* \) does not have any gulps, except possibly at \( t = 0 \) and \( t = T \). A gulp at \( t = T \), however, is clearly sub-optimal since the agent cannot derive any consumption satisfaction from it and the final wealth is also reduced.
All the above considerations lead us to conclude that $C^*$ should take the form of a possible "gulp" at $t = 0$, followed by periods of continuous consumption mixed with periods of no consumption. If we should prescribe a jump $\Delta C^*(0)$ at $t = 0$, then the size of the jump should maximize the value function immediately after the jump. That is, the jump size $\Delta C^*(0)$ should solve the following problem

$$\max_\epsilon J(W(0) - \epsilon, z(0^-) + \beta \epsilon, 0^+).$$

From the first order condition, we then know that $J_W(0^+) = \beta J_z(0^+)$. Moreover, after the initial possible jump at $t = 0$, the following condition should be satisfied for all $t \in (0, T]$:

$$J_W(t) - \beta J_z(t) \geq 0,$$

and at the times when consumption occurs at rates, we must have

$$[J_W(t) - \beta J_z(t)] \epsilon^*(t) = 0.$$

That is, consumption only occurs when $J_W(t) = \beta J_z(t)$.

In the following section, we provide conditions sufficient for a candidate solution to be optimal.

4 Sufficiency

The necessary conditions in (17) and (18) simply say that the agent should adopt a consumption policy of a possible initial "gulp" followed by periods of continuous consumption mixed with periods of no consumption. If the agent starts with a consumption "gulp", then immediately after $t = 0$, the marginal value of wealth $W$ should be equal to ($\beta$ times) the marginal value of past consumption $z$. Otherwise, it would benefit the agent to change the value of the initial "gulp".

During the periods in which the agent consumes continuously, he does this in such a manner that the marginal value of wealth is always equal to ($\beta$ times) the marginal value of average past consumption. In other words, the agent chooses his consumption so that the marginal value of any additional unit of the good at any time is equalized between
the two competing uses available to him, namely investing to increase his wealth and consuming to increase his average past consumption. If this condition is not satisfied in a particular policy, then the agent will find it attractive to deviate from that policy by reallocating the consumption good in the activity with the higher marginal value.

When the agent is not consuming continuously, it must be the case that the marginal utility of wealth is at least as large as ($\beta$ times) the marginal utility of the average past consumption. In other words, during these periods the agent gains more in overall satisfaction by investing all his wealth in the riskless asset and by temporarily refraining from consumption. The increased wealth will provide more consumption and bequeath value in the future than immediate consumption. Of course, when the marginal values of the good in both uses is again equalized, the agent will start consuming again. Finally, the necessary conditions instruct the agent to follow a policy such that at any moment after $t = 0$, the marginal value of the good if consumed, $\beta J_z$, is never strictly greater than the marginal value of the good if invested, $J_W$.

Note that, by the necessary conditions, when consumption occurs, there is no telling whether it is at "rates" or whether it includes singular components. However, the reader will find out later that the closed form solution we construct in the following section for a particular felicity function shows that the optimal consumption policy contains no singular component.

We now provide sufficient conditions for a consumption policy to be optimal. We argue in two steps. First, we show that there is a value function $J^*(W, z, t)$ which is an upper bound on the satisfaction over $[t, T]$ that an agent can obtain starting at $t$ with wealth $W$ and an average past consumption $z$. Second, we give conditions under which a consumption policy achieves this upper bound and hence it is indeed the optimal policy.

Let the agent at any time $t$ have wealth $W(t)$ and average past consumption $z(t^-)$. Assume that the agent has the opportunity to exchange his wealth and average past consumption for a level of satisfaction $J^*(W(t), z(t^-), t)$. In other words, there is a "utility equivalent" function $J^*(W, z, t)$ defined for all levels of wealth and average past consumption and for all times in $[0, T]$ that provides the agent at time $t$ with lump-sum satisfaction for the period $[t, T]$. Therefore the agent has the option of choosing any budget feasible consumption plan till any time $t$, and then exchanging his wealth and
average past consumption at that time for a lump-sum payoff in units of utility. We will show that if \( J^*(W, z, t) \) satisfies certain conditions, then exchanging \( W(0) \) and \( z(0^-) \) at time zero for \( J^*(W(0), z(0^-), 0) \) is better for the agent than adopting any feasible consumption policy till any time \( t \) and then receiving a lump-sum utility at that time.

Consider the situation in which the agent decides to terminate his consumption at time \( t \in [0, T] \), and exchange his state variables at \( t, W(t) \) and \( z(t^-) \), for \( J^*(W(t), z(t^-), t) \). In such a case, the agent’s total satisfaction will be given by:

\[
\int_0^t u(z(s), s) \, ds + J^*(W(t), z(t^-), t). \tag{19}
\]

To ensure that the value in (19) is the correct value of utility for all possible plans, we impose the following boundary conditions on \( J^* \). First, note that the agent might decide not to terminate a given consumption plan at all. At time \( T \), he would receive \( V(W(T)) \) and his total satisfaction will be given by (5). We require that at \( T \),

\[ J^*(W(T), z(T^-), T) = V(W(T)), \]

and thus the value in (19) is the utility for a consumption plan followed completely till time \( T \). Second, note that the agent might follow a consumption plan that exhausts all his wealth strictly before time \( T \), and then “live on” the satisfaction driven from past consumption. To ensure that the value given by (19) is the corresponding utility from such a plan, we impose the boundary condition that

\[ J^*(0, z, t) = \int_t^T u(z e^{-\beta(t-s)}, s) \, ds. \]

The following theorem shows that the agent prefers receiving \( J^*(W(0), z(0^-), 0) \) to the utility gained from any consumption plan terminated at ant time \( t \in (0, T] \).

**Theorem 1** Assume that there exists a differentiable function \( J^*(W, z, t): \mathbb{R}_+ \times \mathbb{R}_+ \times [0, T] \to \mathbb{R} \cup \{-\infty\} \), concave in \( W \) and \( z \), whose partial derivatives are continuous functions of time, that satisfies the following differential equation:

\[
\max \left\{ u(z, t) + J^*_W W r(t) - J^*_z \beta z + J^*_t, \ \beta J^*_z - J^*_W \right\} = 0 \quad \forall t \in [0, T], \tag{20}
\]

together with the boundary conditions:

\[
J^*(W, z, T) = V(W) \quad \text{and} \quad J^*(0, z, t) = \int_t^T u(z e^{-\beta(t-s)}, s) \, ds,
\]

78
\[
\int_0^t u(z(s), s) ds + J^*(W(t), z(t^-), t) \leq J^*(W(0), z(0^-), 0) \quad \forall C \in \mathcal{A}(W(0), 0),
\]

where \(\{z(s); 0 \leq s \leq t^-\}\) and \(W(t)\) are the average past consumption and time \(t\) wealth, respectively, associated with \(C\).

**Proof.** Assume that the agent adopts a consumption plan \(C \in \mathcal{A}(W(0), 0)\) that terminates at \(t\), and let \(s\) points of discontinuity of \(C\) on \([0, t)\) be \(\tau_1, \tau_2, \ldots\). The wealth process and the average past consumption associated with \(C\) is \(\{W(s); 0 \leq s \leq t\}\) and \(\{z(s); 0 \leq z(s) < t\}\), respectively. Recall that \(W\) is left-continuous and \(z\) is right-continuous. For convenience of notation, we at times denote \(J^*(W(t), z(t^-), t), J^*_w(W(t), z(t^-), t), J^*_z(W(t), z(t^-), t),\) and \(J^*_r(W(t), z(t^-), t)\) by \(J^*(t), J^*_w(t), J^*_z(t),\) and \(J^*_r(t)\), respectively. We have

\[
\begin{align*}
\int_0^t u(z(s), s) ds &+ J^*(W(t), z(t^-), t) \\
&= \int_0^t u(z(s), s) ds + J^*(W(0), z(0^-), 0) \\
&\quad + \sum_{i=0}^{\tau_i^+} (J^*_w(s) r W(s) ds - J^*_w(s) d C(s) + J^*_z(s) \beta(d C(s) - z(s) ds) + J^*_r(s) ds) \\
&\quad + \sum_{i=0}^{\tau_i^-} (J^*(\tau_i^+) - J^*(\tau_i^-)),
\end{align*}
\]

(21)

where the equality follows from fundamental theorem of calculus and we have set \(\tau_0 = 0\)

Note that the boundary conditions for \(J^*\) ensure that the left-hand expression is the correct expression of utility even for consumption plans that do not terminate before \(T\), or those plans for which the wealth reaches zero strictly before \(T\).

By the hypothesis that \(J^*\) is concave in its first two arguments, we know that

\[
J^*(\tau_i^+) - J^*(\tau_i^-) \leq J^*_w(\tau_i)(W(\tau_i^+) - W(\tau_i)) + J^*_z(\tau_i)(z(\tau_i) - z(\tau_i^-))
\]

\[
= J^*_w(\tau_i)(C(\tau_i^+) - C(\tau_i)) + J^*_z(\tau_i) \beta(C(\tau_i) - C(\tau_i^-)).
\]

Substituting this into (21) gives

\[
\int_0^t u(z(s), s) ds + J^*(W(t), z(t^-), t) \leq J^*(W(0), z(0^-), 0) + \int_0^t [u(z(s), s) + J^*_w(s) W(s) r(s) - J^*_z(s) \beta z(s) + J^*_r(s)] ds
\]

79
\[ + \int_{0^-}^t [\beta J^*_z(s) - J^*_W(s)] dC(s) \]
\[ \leq J^*(W(0), z(0^-), 0), \]

where the second inequality follows from the hypothesis that the integrands are negative and \( C \) is increasing.

Now, if we can show that there exists an actual policy \((C^*, W^*(T)) \in \mathcal{A}(W(0), 0)\) such that
\[ U(z(0^-), 0; (C^*, W^*(T))) = J^*(W(0), z(0^-), 0), \]

then \((C^*, W^*(T))\) is the optimal policy.

In fact, we will prove more than that. In the following theorem, we provide a characterization of a budget feasible policy \( C^* \), starting from any time \( t \) and any state \( W(t), z(t^-) \), which guarantees that the associated value function is equal to the function \( J^*(W, z, t) \) defined in theorem 1.

**Theorem 2** Suppose that there exists a differentiable function \( J^*(W, z, t): \mathbb{R}_+ \times \mathbb{R}_+ \times [0, T] \to \mathbb{R} \cup \{-\infty\} \), concave in \( W \) and \( z \), whose partial derivatives are continuous functions of time, which solves the differential equation:
\[ \max \{u(z, t) + J^*_W W r(t) - J^*_z \beta z + J^*_t, \beta J^*_z - J^*_W \} = 0 \quad \forall t \in [0, T], \tag{22} \]

Together with the boundary conditions:
\[ J^*(W, z, T) = V(W) \quad \text{and} \]
\[ J^*(0, z, t) = \int_t^T u(z e^{-\rho(s-t)}, s) ds. \]

Starting from any time \( \tau \in [0, T] \), with \( W(\tau) \) and \( z(\tau^-) \), assume that there exists a policy \( C^* \in \mathcal{A}(W(\tau), \tau) \), which is continuous on \((\tau, T]\), with a possible jump \( \Delta C^*(\tau) \), and whose associated state variables are \( W^*(t), z^*(t^-) \), such that for all \( t \in (\tau, T] \):
\[ u(z^*, t) + J^*_W W^* r(t) - J^*_z \beta z^* + J^*_t = 0 \quad \text{and} \]
\[ \int_t^{t+\epsilon} [J^*_W(W^*, z^*, t) - \beta J^*_z(W^*, z^*, t)] dC^*(t) = 0 \quad \text{for all} \quad \epsilon > 0, \tag{24} \]

and
\[ J^*(W(\tau), z(\tau^-), \tau) = J^*(W(\tau) - \Delta C^*(\tau), z(\tau^-) + \beta \Delta C^*(\tau), \tau), \tag{25} \]
then $U(z(t^-), t; (C^*, W^*(T))) = J^*(W(t), z(t^-), t)$ for all $(W(t), z(t^-), t) \in \mathbb{R}_+ \times \mathbb{R}_+ \times [0, T]$. In particular, $U(z(0^-), 0; (C^*, W^*(T))) = J^*(W(0), z(0^-), 0)$. It then follows from theorem 1 that $C^*$ attains the maximum of $U(z(0^-), 0; (C, W(T)))$ over the set $A(W(0), 0)$.

**Proof.** Recall from the definition of $U(z(t^-), t; (C^*, W^*(T)))$ that for $\tau \leq t < s \leq T$, we have:

$$U(z^*(t^-), t; (C^*, W^*(T))) = \int_{t}^{s} u(z^*(\xi), \xi) \, d\xi + U(z^*(s^-), s; (C^*, W^*(T))).$$

(26)

Now consider the following:

$$
\begin{align*}
\int_{t}^{s} u(z^*(\xi), \xi) \, d\xi + J^*(W^*(s), z^*(s^-), s) \\
= \int_{t}^{s} u(z^*(\xi), \xi) \, d\xi + J^*(W^*(t^+), z^*(t), t) + \int_{t}^{s} \left( J_{W}^* \, dW^* + J_{z}^* \, dz^* + J_{\xi}^* \, d\xi \right) \\
= J^*(W^*(t^+), z^*(t), t) + \int_{t}^{s} \left( u(z^*(\xi), \xi) + J_{W}^* \, W^*(\xi) \, r(\xi) - J_{z}^* \, \beta \, z^*(\xi) + J_{\xi}^* \right) \, d\xi \\
+ \int_{t}^{s} \left( J_{W}^* (W^*, z^*, \xi) - J_{W}^* (W^*, z^*, \xi) \right) \, dC^*(\xi) \\
= J^*(W^*(t^+), z^*(t), t) \\
= J^*(W^*(t), z^*(t^-), t),
\end{align*}
$$

(27)

where the first equality follows from the fundamental theorem of calculus. The second equality follows from the dynamics of $W^*$ and $z^*$ and the assumption that $C^*$ is continuous on $(0, T]$, and the third equality follows from equations (23) and (24). The last equality follows from the continuity of $C^*$ for all $t \in (\tau, T]$, and from condition (25), when $t = \tau$.

From equations (26) and (27), we conclude that for all $\tau \leq t < s \leq T$,

$$
U(z^*(t^-), t; (C^*, W^*(T))) - U(z^*(s^-), s; (C^*, W^*(T))) = J^*(W^*(t), z^*(t^-), t) - J^*(W^*(s), z^*(s^-), s) \quad \text{or}
$$

$$U(z^*(t^-), t; (C^*, W^*(T))) - J^*(W^*(t), z^*(t^-), t) = m \quad \text{for all } i \in [\tau, T],
$$

where $m$ is a constant independent of $t$. Now, let $\hat{T} = \min\{t \in [0, T]: W^*(t) = 0\}$, where we have used the convention that if the minimum does not exist, we set the minimum to be $T$. Note from the boundary condition on $J^*$ that: $J^*(W^*(\hat{T}), z^*(\hat{T}^-), \hat{T}) = U(z^*(\hat{T}^-), \hat{T}; (C^*, W^*(T)))$. Thus we conclude that $m = 0$, and the proof is now complete. \qed
5 Optimal Consumption Policy

We provide a closed form solution for the optimal consumption problem formulated in section 2 with infinite horizon in a world of constant interest rate \( r \). The felicity function is \( u(z, t) = \frac{1}{\alpha} e^{-\delta t} z^\alpha \), where \( \alpha < 1 \), and where the discount factor \( \delta \geq 0 \) captures the agent's impatience. In other words, the agent seeks to maximize

\[
U(z(0^-), 0, C) \equiv \int_0^\infty \frac{1}{\alpha} e^{-\delta t} z(t)^\alpha \, dt ,
\]

given \( W(0) \) and \( z(0^-) \) and where \( z(t) \) is given by equation (3). The sufficiency theorem for this problem, which is a modified version of theorem 2, is given in the following corollary, whose proof is omitted.

**Corollary 1** Suppose that there exists a concave, continuously differentiable function \( J^*(W, z) : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R} \cup \{-\infty\} \), which solves the differential equation:

\[
\max\left\{ \frac{z^\alpha}{\alpha} + J^*_W W'r - J^*_z \beta z - \delta J^* \beta J^*_d - J^*_W \right\} = 0 ,
\]

(29)

together with the boundary conditions:

\[
J^*(0, z) = \int_t^\infty e^{-\delta s} \frac{1}{\alpha} (ze^{-\beta(t-t)})^\alpha \, ds \quad \text{and}
\]

\[
\lim_{T \to \infty} e^{-\delta T} J^*(W(T), z(T)) = 0 ,
\]

(30)

(31)

for all feasible policies. Starting from any time \( \tau \in [0, \infty) \), with \( W(\tau) \) and \( z(\tau^-) \), assume that there exists a policy \( C^* \in A(W(\tau), \tau) \), which is continuous, with a possible jump \( \Delta C^*(\tau) \), and whose associated state variables are \( W^*(t), z^*(t^-) \), such that for all \( t \in (\tau, \infty) \):

\[
\frac{z^\alpha}{\alpha} + J^*_W W'r - J^*_z \beta z^* - \delta J^* = 0 \quad \text{and}
\]

\[
\int_t^{t+\epsilon} [J^*_W(W^*, z^*, t) - \beta J^*_z(W^*, z^*, t)] dC^*(t) \quad \text{for all } \epsilon > 0 ,
\]

(32)

(33)

and

\[
J^*(W(\tau), z(\tau^-)) = J^*(W(\tau) - \Delta C^*(\tau), z(\tau^-) + \beta \Delta C^*(\tau)) ,
\]

then \( U(z(t^-), t; C^*) = J^*(W(t), z(t^-)) \) for all \( (W(t), z(t^-)) \in \mathbb{R}_+ \times \mathbb{R}_+ \).

In particular, \( U(z(0^-), 0; C^*) = J^*(W(0), z(0^-)) \). It then follows from theorem 1 that \( C^* \) attains the maximum of \( U(z(0^-), 0; C) \) over the set \( A(W(0), 0) \).
We will show that the key feature of the solution is a critical ratio \( k^* > 0 \) of wealth \( W \) to average past consumption \( z \). If the agent starts at time zero with \( \frac{W(0)}{z(0)} \) strictly less than \( k^* \), then the optimal behavior is to invest all the wealth in the riskless asset and wait while \( W \) increases and \( z \) declines till the ratio \( \frac{W}{z} \) reaches \( k^* \). From then on, the agent consumes at the rate which keeps the ratio \( \frac{W}{z} \) equal to \( k^* \) forever. If, on the other hand, the agent starts with \( \frac{W(0)}{z(0)} \) strictly greater than \( k^* \), then the optimal behavior is to take a "gulp" of consumption, reducing \( W \) and increasing \( z \), to bring \( \frac{W}{z} \) immediately to \( k^* \). Following this gulp, the optimal consumption occurs at the rate that keeps the ratio \( \frac{W}{z} \) equal to \( k^* \) forever. The critical ratio \( k^* \) depends on the interest rate \( r \), the impatience of the agent captured by \( \delta \), the concavity of the felicity function captured by \( \alpha \), and the rate of decay of past consumption captured by \( \beta \).

Our construction of the critical ratio \( k^* \) and hence the optimal solution is in two steps. First, we examine candidate solutions of the \( k \)-ratio form. For any \( k > 0 \), the \( k \)-ratio policy is the policy of keeping \( \frac{W}{z} \) equal to \( k \) forever, after an initial "gulp" if \( \frac{W(0)}{z(0)} > k \), or after a period of no consumption if \( \frac{W(0)}{z(0)} < k \). We will show that the value function associated with any \( k \)-ratio policy, denoted \( J^k(W, z) \), satisfies:

\[
\frac{z^\alpha}{\alpha} + rWJ^k_W - \beta zJ^k_z - \delta J^k = 0 \quad \text{if} \quad \frac{W}{z} \leq k \quad \text{and} \quad J^k_W - \beta J^k_z = 0 \quad \text{if} \quad \frac{W}{z} \geq k. \]

Second, we show that there is a unique value \( k^* > 0 \) such that the associated value function, \( J^*(W, z) \), satisfies the differential inequality

\[
\max\left\{ \frac{z^\alpha}{\alpha} + J^*_W rW - J^*_z \beta z - \delta J^*, \beta J^*_z - J^*_W \right\} = 0, \tag{35}
\]

together with the boundary condition that

\[
J(0, z) = \frac{1}{\alpha} \int_0^\infty e^{-\lambda t} \left(ze^{-\beta t}\right)^\alpha \, dt = \frac{1}{\alpha(\alpha \beta + \delta)} z^\alpha. \tag{36}
\]

It then follows from theorem 1, that the ratio policy associated with \( k^* \) is the optimal solution for the agent's problem. We record our solution in the following statements.
Assumption 1 The parameters of the problem satisfy

\[ \delta > \alpha r \quad \text{and} \quad (1 - \alpha) \beta > \delta - r. \]

This assumption ensures the existence of a solution to the infinite horizon program. It rules out the case when the agent is so patient that he prefers to wait and accumulate wealth for the longest possible time. This behavior leads to infinite utility. Now fix a k-ratio policy.

Lemma 1 Suppose that \( \frac{W(0)}{z(0^-)} = k \). The consumption rate required to keep \( \frac{W(t)}{z(t)} = k \) for all \( t \geq 0 \) is

\[ c(t) = \frac{(r + \beta)}{(1 + \beta k)} W(t), \]

and the corresponding utility is

\[ \frac{1}{\alpha \left[ \delta - \alpha \beta \left( \frac{r k - 1}{\beta k + 1} \right) \right]} z(0^-)^{\alpha}. \]

PROOF. Using the dynamics of \( W \) and \( z \), and equating \( \frac{d}{dt} \log \left( \frac{W}{z} \right) \) to zero, we can compute \( c \). Another simple computation produces the associated utility.

Lemma 2 Suppose that \( \frac{W(0)}{z(0^-)} < k \). Consider the policy of no consumption till \( t^* \) when \( \frac{W(t^*)}{z(t^*)} = k \), and then consuming to keep the ratio \( \frac{W(t)}{z(t)} = k \), for \( t > t^* \). The total utility from this policy is

\[ \frac{z(0^-)^{\alpha}}{\alpha (\delta + \alpha \beta)} + \frac{z(0^-)^{\alpha}}{\alpha \left[ \frac{W(0)}{k z(0^-)} \right]^{\frac{t^* + \alpha \beta}{1 + \beta}} \left\{ \frac{1}{\alpha \beta \left( \frac{r k - 1}{\beta k + 1} \right)} - \frac{1}{\delta + \alpha \beta} \right\}. \] (37)

PROOF. The reader can easily verify that \( t^* = \log \left[ \frac{k z(0^-)}{W(0)} \right]^{\frac{1}{r + \beta}} \). A direct computation establishes the lemma.

Proposition 1 The value function \( J^k(W, z) \) associated with any k-ratio policy is given by:

\[ J^k(W, z) = \begin{cases} \frac{z^\alpha}{\alpha (\delta + \alpha \beta)} + \frac{z^\alpha}{\alpha \left( \frac{W}{k z} \right)^{\frac{t^* + \alpha \beta}{1 + \beta}}} A & \text{if } \frac{W}{z} \leq k \\ \frac{1}{\alpha \left( \frac{t^* + \alpha \beta W}{1 + \beta k} \right)^{\alpha}} B & \text{if } \frac{W}{z} \geq k \end{cases} \]
\[ A = \frac{1}{\delta - \alpha \beta (\frac{r}{\beta} + 1)} - \frac{1}{\delta + \alpha \beta} \quad \text{and} \]
\[ B = \frac{1}{\delta - \alpha \beta (\frac{r}{\beta} + 1)} \].

Furthermore, \( J^k \) is continuous, concave, has continuous first derivatives and satisfies
\[ \frac{z^\alpha}{\alpha} + r W J^k_W - \beta z J^k_z - \delta J^k = 0 \quad \text{if} \quad \frac{W}{z} \leq k, \]
\[ J^k_W - \beta J^k_z = 0 \quad \text{if} \quad \frac{W}{z} \geq k, \quad (38) \]
together with the boundary condition
\[ J(0, z) = \frac{1}{\alpha (\alpha \beta + \delta)} z^\alpha. \]

PROOF. Suppose that \( \frac{W}{z} > k \). The size of the initial "gulp" required to bring the ratio immediately to \( k \) is
\[ \Delta = \frac{W - kz}{1 + \beta k}, \]
and we define \( J^k(W, z) \equiv J^k(W - \Delta, z + \beta \Delta) \). Concavity of \( J^k \) in both \( W \) and \( z \) follows from assumption 1, and the rest of the proposition can be easily verified by direct computations. \( \blacksquare \)

This proposition shows that any ratio policy has a value function which satisfies the differential equations (38) in the relevant parts of the domain. The value function of the optimal solution, however, satisfies the differential inequality (35) over the whole domain. There is exactly one value \( k^* \) that produces the optimal policy.

**Proposition 2** Let
\[ k^* = \frac{\frac{r - \delta}{\beta} + (1 - \alpha)}{\delta - \alpha \beta}, \]
and note that \( k^* > 0 \) by assumption 1. The value function \( J^* \) associated with the \( k^* \)-ratio policy is continuous, concave, has continuous first derivatives and satisfies the differential inequality:
\[ \max \left\{ \frac{z^\alpha}{\alpha} + J^*_W r W - J^*_z \beta z - \delta J^* - \beta J^*_z - J^*_W \right\} = 0, \]
85
together with the boundary condition that

$$J(0, z) = \frac{1}{\alpha(\alpha \beta + \delta)} z^\alpha.$$ 

It then follows from theorem 1, that the $k^*$-ratio policy is the optimal solution for the agent's problem.

**Proof.** In light of proposition (1), we only need to prove that

$$\frac{z^\alpha}{\alpha} + J_W^* r W - J_z^* \beta z - \delta J^* \leq 0 \quad \text{if} \quad \frac{W}{z} \geq k^* \quad \text{and} \quad \beta J_z^* - J_w^* \leq 0 \quad \text{if} \quad \frac{W}{z} \leq k^*.$$ 

To prove the first inequality, consider any point $\bar{z} \equiv (W, z)$ such that $W \geq k^* z$. Referring to Figure 1, let $\bar{a}$ be the point on the intersection of the line $W = k^* z$ and the straight line passing through the point $\bar{z}$ with slope $\frac{dW}{dz} = -\frac{1}{\beta}$. In other words, $\bar{a}$ is the point on the boundary $W = k^* z$, to which one would jump if one starts at $\bar{z}$. By construction, $J^*(\bar{z}) = J^*(\bar{a})$.

Let

$$f(W, z) = \frac{z^\alpha}{\alpha} + J_W^* r W - J_z^* \beta z,$$

and note that $f(\bar{a}) - \delta J^*(\bar{a}) = 0$. Applying the fundamental theorem of calculus along the straight line connecting $\bar{a}$ and $\bar{z}$, we obtain

$$f(\bar{z}) - f(\bar{a}) = \int_{\bar{a}}^{\bar{z}} f_w dW + f_z dz.$$ 

Noting that along the line connecting $\bar{a}$ and $\bar{z}$, we have $dz = -\beta dz$, and that $dW > 0$ in the direction from $\bar{a}$ to $\bar{z}$, it then follows that $f_w - \beta f_z \leq 0$ is sufficient to conclude that

$$\frac{z^\alpha}{\alpha} + J_W^* r W - J_z^* \beta z - \delta J^* \leq 0 \quad \text{if} \quad \frac{W}{z} \geq k^*.$$ 

Computing $f_w - \beta f_z$ in the region $W \geq k^* z$, the reader can easily verify that $f_w - \beta f_z \leq 0$ for $k^* = \frac{r \beta + (1 - \alpha)}{r + \alpha r}$.

To prove the second inequality, consider the function $J_W^* - \beta J_z^*$ in the region where $W \leq k^* z$. We can write

$$J_W^* - \beta J_z^* = z^{\alpha-1} g\left(\frac{k^* z}{W}\right) \quad \text{where} \quad g(y) = A\left[\frac{y}{k^*} \left(\frac{\delta + \alpha \beta}{r + \beta}\right) + \frac{\beta(\delta - \alpha r)}{r + \beta}\right] y^{-\frac{\delta + \alpha \beta}{r + \beta}} - \left[\frac{\beta}{\alpha \beta + \delta}\right].$$

86
where \( A \) is given in proposition (1). Note that \( g(1) = 0 \), and that \( 1 - \frac{\delta + \alpha \beta}{r + \delta} > 0 \), by assumption 1, hence \( g(y) \uparrow \infty \) as \( y \uparrow \infty \).

Computing the derivative of \( g \), we get

\[
\frac{d g}{d y} = \frac{A(\delta + \alpha \beta)}{r + \beta} y^{-\frac{\alpha r \beta}{r + \beta}} \left[ \frac{1}{k^*} - \frac{\delta + \alpha \beta}{k^*(r + \beta)} - \frac{\beta \delta - \alpha r}{y(\beta + r)} \right].
\]

Substituting \( k^* \) in the expression for \( \frac{d g}{d y} \), we conclude that \( \frac{d g}{d y} = 0 \) for \( y = 1 \), and that \( \frac{d g}{d y} > 0 \) for \( y > 1 \). It then follows that \( J_W^* - \beta J_\gamma^* \geq 0 \), for all points \((W, z)\) such that \( W \leq k^* z \).

We have thus shown that the nature of the optimal consumption policy is exerting control over the ratio of wealth to average past consumption to keep it at or below a critical level. Any policy of keeping the ratio at or below a fixed constant satisfies many of the sufficient conditions for optimality. However, the differential inequality (35), which is the crucial condition, is satisfied by only one critical ratio \( k^* \).

6 Concluding Remarks

In this essay, we have provided sufficient conditions for a consumption policy to be optimal for a class of time-nonseparable preferences that treat consumptions at nearly adjacent dates to be almost perfect substitutes. We demonstrated our general theory by explicitly solving in closed form the optimal consumption policy for a particular felicity function.

The heuristically derived necessary conditions for optimality in Section 3 can be justified rigorously. We refer the reader to Blaquière (1985) for details.

We are currently working on incorporating the notion of habit formation into our model while preserving the feature that consumptions at nearby dates are almost perfect substitutes. We hope to report our findings in the future.

7 References

1. K. Arrow and M. Kurz, Public Investment, the Rate of Return, and Optimal Fiscal Policy, Johns Hopkins Press, Baltimore, 1970.


The State Space showing the Boundary B. For all times after \( t=0 \), the optimal policy restricts the state to the admissible region \( O \).

**Figure 1**
Trajectory of wealth along the optimal path for time-additive and non-time-additive preferences assuming

$$\delta = r, \beta_1 > \beta_2$$

**Figure 2**
Essay III

OPTIMAL CONSUMPTION WITH INTERTEMPORAL SUBSTITUTION II:
THE CASE OF UNCERTAINTY

Abstract

We study the problem of optimal consumption and portfolio choice in continuous time under uncertainty for a class of utility functions that capture the notion that consumptions at nearby dates are almost perfect substitutes. The class we consider excludes all time-additive and almost all the non-time-additive utility functions used in the literature. We provide sufficient conditions for a consumption and portfolio policy to be optimal. Furthermore, we demonstrate our general theory by solving in a closed form the optimal consumption and portfolio policy for a particular felicity function when the prices of the assets follow a geometric Brownian motion process. The optimal consumption policy in our solution consists of a possible initial "gulp" of consumption followed by a consumption and investment behavior in which consumption occurs only periodically. The agent controls the ratio of wealth to average past consumption and uses current consumption to ensure that this ratio does not exceed a critical level.
1 Introduction and Summary

This is the second part in a series of two essays. In the first part, Essay I of this dissertation, we study the problem of optimal consumption choice in continuous time under certainty for a class of agents who treat consumptions at nearby dates as almost perfect substitutes. In this essay, we study the problem of optimal intertemporal consumption and portfolio rules under uncertainty for the same class of agents.

The seminal work in the area of optimal consumption and portfolio choice in continuous time under uncertainty is Merton (1971). The agents in Merton's model, however, have time additive utility functions. This implies that the preferences of these agents do not exhibit the intuitively appealing notion that past consumption can contribute to the satisfaction in the future. Many researchers have recently used non-time-additive utilities; see, for example, Epstein (1987) who studies the utility form introduced by Koopmans (1960) and Uzawa (1968) in the context of continuous time equilibrium models, and Constantinides (1988), Heaton (1990), and Sundaresan (1989) who study habit formation models.

Huang and Kreps (1989), in a model of certainty, and Hindy and Huang (1989a), in a model of uncertainty, study preferences which exhibit the notion that consumption at nearby dates are almost perfect substitutes. They show, among other things, that most of the currently used non-time-additive utilities fail to capture this notion because the index of instantaneous satisfaction, or felicity, at any time depends explicitly on the consumption rate at that time. The class of utility functions that we consider in this essay has the key feature that the felicity function at any time depends only upon an exponentially weighted average of past consumption – one derives satisfaction only from past consumption. This feature is the main difference between the utility function we study here and those studied by other researchers, except for Heaton (1990). We will demonstrate the tractability of this class of utility by showing how dynamic programming can be used to give sufficient conditions for optimality. In particular, for a specific functional form of the felicity function, we solve in closed form the optimal consumption policy when asset prices follow a geometric Brownian motion.

In Essay II, we consider an economic agent who lives from time $t = 0$ to $t = T$ in a
certain world where there is a single consumption good available at any time between 0 and $T$. We represent the agent's consumption pattern over his life span by a positive\(^1\), increasing, right continuous function $C:[0,T] \rightarrow \mathbb{R}_+$, with $C(t)$ denoting the cumulative consumption from time zero to time $t$. For a consumption pattern $C$, we define the exponentially weighted average past consumption, $z$, by:

$$z(t) = z(0^-)e^{-\beta t} + \beta \int_{0^-}^{t} e^{-\beta(t-s)} dC(s),$$

(1)

where $z(0^-) \geq 0$ is a constant and $\beta$ is a weighting factor. Given $z(0^-)$, the agent's utility for the consumption pattern $C$ and final wealth $W(T)$ is given by:

$$U(z(0^-),0;(C,W(T))) = \int_{0}^{T} u(z(t),t) dt + V(W(T)),$$

(2)

where both $u$ and $V$ are continuous and increasing functions. Furthermore, $V$ is strictly concave and $u$ is strictly concave in its first argument.

The agent starts at time 0 with endowment $W(0)$, and at any time $t$ he can invest part of or all his wealth in a riskless asset with instantaneous rate of return $r(t)$. Starting from a given level of wealth, each consumption plan $C$ determines for every time $t$, the wealth $W(t)$, and the average past consumption $z(t^-)$. Note that at any $t$, the value $z(t^-)$ is the average past consumption up till $t$, excluding the possible consumption at $t$. We will call $W(t)$ and $z(t^-)$ the state variables at $t$, since they represent the status of the agent before making his consumption decision at $t$.

In Essay II, we provide sufficient conditions for a consumption policy $C^*$ to be optimal. These conditions are, roughly, that for all times after $t = 0$, the marginal value of wealth is at least as large as ($\beta$ times) the marginal value of average past consumption and that consumption occurs only when these marginal values are equal.

Using the sufficient conditions, we construct an explicit solution for specific felicity and bequeath functions in an infinite horizon program. The basic insight behind our construction is that the object of control is the ratio of $W$ to $z$. We show that there is a critical ratio, $k^*$, of wealth to average past consumption that should not be exceeded at any moment. If, in the absence of consumption, the agent's ratio of wealth to his average

---

\(^1\) We use weak relations in our discussion. Hence increasing is equivalent to nondecreasing, for example. When a relation is strict, we will explicitly state so.
past consumption is strictly less than $k^*$, he will choose to refrain from consumption and invest all of his wealth in the riskless asset. If, on the other hand, $\frac{W}{z}$ is strictly higher than $k^*$, then the agent will choose to consume a "gulp" to reduce $\frac{W}{z}$ to the critical ratio. Finally, the agent consumes at rates only when $\frac{W}{z}$ is equal to $k^*$, and consumption occurs at the minimum rate required to keep $\frac{W}{z}$ from exceeding the critical ratio. These elements of our construction lead to an optimal solution with an initial possible "gulp" followed by periods of consumption at rates mixed with periods of no consumption.

In this essay, we consider an agent in an environment of uncertainty and provide sufficient conditions for a consumption plan and portfolio policy to maximize his preferences. The sufficient conditions have the same flavor as in the certainty case. It is sufficient that the value function, or the maximum attainable utility starting from any state, satisfies a differential inequality which can be viewed as an application of the Bellman optimality principle. In particular, the conditions require that at all times the marginal value of wealth is at least as large as ($\beta$ times) the marginal value of the average past consumption. Furthermore, consumption should only occur when these marginal values are equal.

We also construct a closed form solution of the optimal consumption problem for a particular class of felicity and bequest functions in the case when the prices of the risky assets follow a geometric Brownian motion. The idea of the optimal solution is the same as in the certainty case: keep the ratio $\frac{W}{z}$ less than a critical number, say $k^*$, and consume only when $\frac{W}{z} = k^*$. In the case of uncertainty, however, a new and totally distinct phenomenon appears. The ratio $\frac{W}{z}$ is subject to random shocks because of the randomness of the return on the risky assets. Furthermore, the sample path properties of the Brownian motion lead to peculiar behavior of the quantity $\frac{W}{z}$ during the times when it is very close to the critical number $k^*$. The quantity $\frac{W}{z}$ fluctuates so fast that whenever it hits the value $k^*$ it bounces back and forth to hit it again uncountably infinite number of times. This results in a process of optimal cumulative consumption with nontrivial increasing sample paths in which consumption occurs only periodically.

The rest of this essay is organized as follows. Section 2 sets up the consumption and portfolio problem under uncertainty in continuous time. Section 3 provides sufficient conditions for optimality. In Section 4, we solve the optimal consumption and portfolio
problem in closed form for a particular felicity function when the prices of risky assets follow a process of geometrical Brownian motion. Section 5 contains concluding remarks.

2 Formulation

Consider an economic agent who lives from time $t = 0$ to $t = T$ in a world of uncertainty where there is a single consumption good available at any time between 0 and $T$. The nature of uncertainty and the opportunities for consumption and investment available to the agent are described in the following.

2.1 Uncertainty and Information Resolution

The primitive source of uncertainty in the agent’s world is modeled by a complete probability space $(\Omega, \mathcal{F}, P)$. Over the agent’s life span $[0, T]$, there is an $M$-dimensional standard Brownian Motion defined on the probability space and denoted by $B = \{B_m(t); t \in [0, T], m = 1, 2, \ldots, M\}$. Information is revealed to the agent via the realizations of this Brownian Motion. We model this structure of information resolution by $\mathcal{F} = \{\mathcal{F}_t; t \in [0, T]\}$, the family of increasing sub-sigma-fields of $\mathcal{F}$, or the filtration, generated by $B$.

We assume that $\mathcal{F}$ is complete by assuming that $\mathcal{F}_0$ contains all the $P$-null subsets and we note that $\mathcal{F}$ is right-continuous. We also assume that all uncertainty is resolved by time $T$, or $\mathcal{F}_T = \mathcal{F}$. On the other hand, $\mathcal{F}_0$ is almost trivial since for an $M$-dimensional standard Brownian Motion, $B(0) = 0$ a.s.$^2$

A process $Y$ is a mapping $Y: \Omega \times [0, T] \to \mathbb{R}$ that is measurable with respect to $\mathcal{F} \otimes B([0, T])$, the product sigma-field generated by $\mathcal{F}$ and the Borel sigma-field of $[0, T]$. For each $\omega \in \Omega$, $Y(\omega, .): [0, T] \to \mathbb{R}$ is a sample path and for each $t \in [0, T], Y(., t): \Omega \to \mathbb{R}$ is a random variable. The process $Y$ is said to be adapted to $\mathcal{F}$ if for each $t \in [0, T]$, $Y(t)$ is $\mathcal{F}_t$-measurable. This is a natural information constraint: the value of the process at time $t$ cannot depend on information yet to be revealed. For brevity, all processes to be discussed will be adapted to $\mathcal{F}$. Finally, we introduce the notion of optional time by recording the following definition:

$^2$We use the notation a.s. to denote statements which are true with probability one.
Definition 1 (Optional Time) The function \( g: \Omega \to [0, \infty] \) is an optional time with respect to \( F \) if

\[
\{ \omega \in \Omega: g(\omega) \leq t \} \in F_t \quad \forall t \in [0, T] .
\]

An optional time can always be interpreted as the first time a particular event happens. The condition \( \{ \omega \in \Omega: g(\omega) \leq t \} \in F_t \) in the above definition then says that at any time \( t \), it will be known whether that particular event has happened or not.

2.2 Consumption Space

The agent can consume the good at "gulps" at any moment, and can consume at finite rates over intervals. He can also refrain from consumption altogether for some time. Moreover, the sample path of his accumulated consumption at any time \( t \) can have a singular component, that is a continuous nontrivial increasing function whose derivative is zero for almost all \( t \), \( P - a.s. \).

Let \( X_+ \) be the space of all processes \( x \) whose sample paths are positive, increasing and right continuous. The linear span of \( X_+ \), which is denoted by \( X \), is the space of all processes with right continuous sample paths of finite variation. Recall that a finite variation function \( x(\omega,.) \) on \([0, T]\) has a finite left-limit at any \( t \in (0, T] \) denoted by \( x(\omega, t^-) \) and a finite right-limit at any \( t \in [0, T) \) denoted by \( x(\omega, t^+) \). By right continuity of sample paths, for any process \( x \in X \) we have \( x(\omega, t^+) = x(\omega, t) \), \( P - a.s., t \in (0, T) \). We will use the convention that \( x(\omega, 0^-) = 0 \), \( P - a.s. \). Since left limits exist for the sample paths of any \( x \in X \), a jump of \( x(\omega, .) \) at \( \tau \) is \( \Delta x(\omega, \tau) \equiv x(\omega, \tau) - x(\omega, \tau^-) \).

Let \( p \geq 1 \) be fixed. For any \( x \in X \), define the functional:

\[
\|x\| = \left( \mathbb{E} \left[ \int_0^T |x(t)|^p + |x(T)|^p \right] \right)^{\frac{1}{p}} .
\]  

(3)

The consumption set available to the agent, \( \mathcal{L} \), is all those consumption patterns \( C \) for which the above functional is finite, or

\[
\mathcal{L} = \{ C \in X_+: \|C\| < \infty \} ,
\]

and the functional defined in (3) defines a norm on \( \mathcal{L} \).
The stochastic process \( C \in \mathcal{L} \) is a consumption pattern available to the agent with \( C(\omega, t) \) denoting the accumulated consumption from time 0 to time \( t \) in state \( \omega \). For any \( \omega \in \Omega \), the points of discontinuity of \( C(\omega, t) \) are the moments when the agent consumes a "gulp". Moreover, \( C(\omega, t) \) has an absolutely continuous component over the intervals during which the agent is consuming at rates \( c(\omega, t) \). Finally, \( C(\omega, t) \) may have a singular part. Singular components of consumption processes will play an important role in our analysis of the optimal consumption policy.

Note that the norm defined in equation (3) is a standard \( L^p \) norm. However, here it is defined on cumulative consumption rather than on consumption rates. This norm defines a topology on the set \( \{ x \in X : \| x \| < \infty \} \), which we denote by \( \mathcal{T} \). Hindy and Huang (1989a) study the properties of a family of topologies that include \( \mathcal{T} \) and show that preferences continuous in \( \mathcal{T} \), relativized to \( \mathcal{L} \), treat consumption at nearby dates as almost perfect substitutes. We will consider agents whose preferences are continuous in \( \mathcal{T} \).

2.3 Investment Opportunities

The agent has the opportunity to invest his wealth in a frictionless securities market with \( N + 1, N \leq M \), long lived securities, continuously traded, and indexed by \( n = 0, 1, 2, \ldots, N \). Security \( n \), where \( n = 1, 2, \ldots, N \), is risky, pays dividends at rate \( \rho_n(t) \), and sells ex-dividend for \( S_n(t) \), at time \( t \). We assume that \( \rho_n(t) \) can be written as \( \rho_n(S(t), t) \) with \( \rho_n(y, t) : \mathbb{R}^N \times [0, T] \rightarrow \mathbb{R} \) Borel measurable, and we will use \( S(t) \) to denote the column vector \( [S_1(t), S_2(t), \ldots, S_N(t)]^T \). The agent also has the opportunity to borrow or lend funds at the instantaneous riskless rate \( r(t) \), which we will assume can be written as \( r(t) = r(S(t), t) \).

The price process for the risky securities follows an Itô process given by:

\[
S(t) + \int_0^t \rho(S(s), s) \, ds = S(0) + \int_0^t \mu(S(s), s) \, ds + \int_0^t \sigma(S(s),s) \, dB(s) \quad \forall t \in [0, T] \quad \text{a.s.,}
\]

where \( \mu(y, t) - \rho(y, t) : \mathbb{R}^N \times [0, T] \rightarrow \mathbb{R}^N \) and \( \sigma(y, t) : \mathbb{R}^N \times [0, T] \rightarrow \mathbb{R}^{N \times M} \) are continuous in \( y \) and \( t \). We impose the following conditions on the parameters \( \mu(y, t) - \rho(y, t) \) and

\[\text{The superscript } ^T \text{ denotes transpose.}\]
\[ \sigma(y,t) \] to ensure existence and uniqueness of solution to equation (4), see, for example, Fleming and Rishel (1975, theorem V.4.1).

If \( \mu \) is a vector in \( \mathbb{R}^N \), let \( |\mu| \) be the Euclidean norm of \( \mu \). In addition, if \( \sigma \) is a matrix, let \( |\sigma|^2 \) denote \( \text{tr}(\sigma \sigma^T) \), where \( \text{tr} \) is the trace of a square matrix.

**Assumption 1 (Uniform Growth and Local Lipschitz Conditions)** There exist a strictly positive constant \( K_1 \) such that for all \( t \),

\[ |\mu(y,t) - \rho(y,t)| \leq K_1 (1 + |y|), \quad |\sigma(y,t)| \leq K_1 (1 + |y|) \quad \forall y \in \mathbb{R}^N. \tag{5} \]

Furthermore, for any bounded \( \Theta \subset \mathbb{R}^N \) and for any \( t \in [0,T] \), there exists a strictly positive constant \( K_2 \), which may depend on \( \Theta \) and \( t \), such that:

\[ |(\mu(y_1,t) - \rho(y_1,t)) - (\mu(y_2,t) - \rho(y_2,t))| \leq K_2 |y_1 - y_2|, \quad |\sigma(y_1,t) - \sigma(y_2,t)| \leq K_2 |y_1 - y_2| \quad \forall y_1, y_2 \in \Theta. \tag{6} \]

The agent manages his wealth dynamically by choosing a consumption plan \( C \in \mathcal{L} \) and an investment policy \( A = [A_1, A_2, \ldots, A_N]^T \), an \( N \)-dimensional stochastic process, where \( A_n(t) \) is the proportion of wealth invested in the \( n \)-th risky asset, before trading and consuming at \( t \). In order to keep his budget balanced at all times, the agent lends or borrows the excess or shortage of funds at the instantaneous interest rate. We will specify the space of admissible trading strategies precisely in section 3. We note here that they will be a subset of the trading strategies for which the stochastic integral in the budget constraint, to be given shortly, is well defined.

For a given consumption and investment policy, the agent's wealth evolves according to the following dynamics:

\[
W(t) = W(0) + \int_0^t \left( W(s)\rho(s) + W(s)A^T(s)I_{S-1}(s)(\mu(s) - r(s)S(s)) \right) ds - C(t^-) + \int_0^t W(s)A^T(s)I_{S-1}(s)\sigma(s) dB(s) \quad \forall t \in [0,T] \quad a.s. \tag{7}
\]

We interpret \( W(t) \) to be the wealth that the agent has at time \( t \) before trading and consumption take place. In addition, we impose the condition that \( W(t) \geq 0 \forall t \in [0,T], P - a.s. \) Note that the wealth process has a jump when the consumption policy calls for a jump, and that it has left continuous sample paths. Finally, we require that the final wealth \( W(T^+) \in L^p_+(\Omega, \mathcal{F}, P) \).
2.4 Preferences

The agent has preferences over life time consumption and final wealth. He treats consumptions at nearby dates as almost perfect substitutes. Delaying or advancing his consumption for a very small period of time, without violating the condition that consumption at \( t \) depends only on information available up to \( t \), has a very small effect on his total satisfaction. The agent’s preferences are given by a utility function from a family studied in Hindy and Huang (1989a).

Fix any consumption pattern \( C \). At any time \( t \), define the process of exponentially weighted average past consumption, \( z \), by:

\[
z(t) = z(0^-)e^{-\beta t} + \beta \int_{0^-}^{t} e^{-\beta(t-s)} dC(s) \quad F - a.s.,
\]

where \( z(0^-) \geq 0 \) is a constant, where \( \beta \) is a weighting factor and where the above integral is defined path by path in the Lebesgue-Stieltjes sense. Note that the lower limit of the integral in (8) is \( 0^- \), to account for the possible jump of \( C \) at \( t = 0 \), and that \( z \) is a right continuous process which jumps whenever \( C \) jumps. Moreover, \( z \) has a singular component whenever \( C(\omega, t) \) does. Also observe that higher values of \( \beta \) imply higher emphasis on the recent past and less emphasis on consumption in the distant past.

Starting at time \( t = 0 \) with wealth \( W(0) \), and given a level of average past consumption \( z(0^-) \), the agent’s utility for the consumption pattern \( C \), and final wealth \( W(T) \) produced using an admissible investment strategy \( A \) is given by:

\[
U(z(0^-), t; (C, A, W(T))) = E\left[\int_{0}^{T} u(z(s), s) dt + V(W(T))|\mathcal{F}_t\right],
\]

where \( u: \mathbb{R}_+ \times [0, T] \rightarrow \mathbb{R} \cup \{-\infty\} \), the felicity function, is increasing and strictly concave in its first argument, and where \( V: \mathbb{R}_+ \rightarrow \mathbb{R} \cup \{-\infty\} \), the bequeath function, is increasing and strictly concave. Note that both functions are state independent. Note also that we have allowed the possibility that \( u \) or \( V \) take the value \(-\infty\) at zero. For example \( V(W) = \frac{W^\alpha}{\alpha} \), for \( \alpha < 1 \). Preferences given by (9) are continuous in the product topology generated by \( T \) and the \( L^p \) topology on \( L^p_+(\Omega, \mathcal{F}, P) \), see Hindy and Huang (1989a), proposition 17.

---

*More precisely, we should write final wealth as \( W(T^+) \). We elect to write it as \( W(T) \) for simplicity of notation. This is rationalized by the observation that a gulp of consumption at \( T \) adds nothing to the agent’s direct satisfaction and reduces the terminal wealth passed on to the next “generation”.*
2.5 Agent’s problem

Starting at any optional time $\tau$ with a level of wealth $W(\tau)$ and average past consumption $z(\tau^-)$, each consumption plan $C$ and portfolio policy $A$ determine the value of final wealth $W(T)$. They also determine the evolution of the state variables $W(s)$ and $z(s^-)$, for $s > \tau$. We call $W(\tau), z(\tau^-)$ and $S(\tau)$ the state variables at time $\tau$. Note that $W(\tau)$ and $z(\tau^-)$ are the value of wealth and average past consumption at $\tau$ before the consumption decision is made. Also note that different prices of the risky assets at $\tau$ represent different states since different prices imply different investment opportunities in the future.

The agent faces the problem of choosing a consumption pattern $C^*$ together with an admissible investment policy $A^*$—the optimal controls—to maximize his utility in (9) subject to the dynamics given in (7) and (8), and given that his initial wealth is $W(0)$ and given $z(0^-)$.

Formally, let $\mathcal{A}(W(0), 0)$ be the space of all tuples of consumption pattern $C$ and admissible investment policy $A$ that satisfy the budget constraint in (7) such that $W(t) \geq 0 \forall t \in [0, T], P - a.s.$ The agent’s problem is to solve the following program:

$$\sup_{(C, A) \in \mathcal{A}(W(0), 0)} U(z(0^-), 0; (C, A, W(T))). \quad (10)$$

A solution to (10) exists if the supremum is finite and is attained by some $(C^*, A^*) \in \mathcal{A}(W(0), 0)$.

At times, we will also consider a sub-problem of (10) at some optional $\tau \in [0, T)$: given $S(\tau), \{C(s); s \in [0, \tau^-]\}$ and $W(\tau)$ determined by (7) (the budget constraint), solve

$$\sup_{(C^{\tau}, A^{\tau}) \in \mathcal{A}(W(\tau), \tau)} U(z(\tau^-), \tau; (C^{\tau}, A^{\tau}, W(T))) \equiv E\left[\int_{\tau}^{T} u(z(s), s)ds + V(W(T))|\mathcal{F}_\tau\right], \quad (11)$$

where $C^{\tau}$ is an increasing and right-continuous process on the stochastic interval $[\tau, T]$ representing cumulative consumption starting from $\tau$ with the convention that $C^{\tau}(s) = 0$ for all $s < \tau$, and where $A^{\tau}$ is an $N$-dimensional process on the stochastic interval $[\tau, T]$, denoting the proportion of wealth invested in the risky assets. $\mathcal{A}(W(\tau), \tau)$ denotes the consumption patterns $C^{\tau}$ and admissible investment policies $A^{\tau}$ that satisfy the budget constraint (7) on $[\tau, T]$ with an initial value $W(\tau)$, and the process $z$ on $[\tau, T]$ is given.

---

6The space $\mathcal{A}(W(0), 0)$ is specified precisely in theorem 1.
by:
\[ z(s) = z(\tau^-)e^{-\beta(s-\tau)} + \int_{\tau^-}^{s} e^{-\beta(s-\xi)}dC^\tau(\xi) \quad \text{for } s \geq \tau \quad P - a.s. \quad (12) \]

It is clear that once \( S(\tau), W(\tau) \) and \( z(\tau^-) \) are known, \( \{C(s); s \in [0, \tau^-]\} \) has no impact on the choices of \( C^\tau \) or \( A^\tau \) in (11) and this justifies the notation \( U(z(\tau^-), \tau; (C^\tau, A^\tau, W(T))) \). Moreover, if \( (C^*, A^*) \) is a solution to (10), then \( \{C^{**}(s) \equiv C^*(s) - C^*(\tau^-); s \in [\tau, T]\} \) and \( A^{**} \), which is \( A^* \) restricted to the interval \( [\tau, T] \), is a solution to (11) with \( z(\tau^-) \) and \( W(\tau) \) corresponding to \( (C^*, A^*) \).

3 Sufficiency

In this section we provide sufficient conditions for a candidate policy to be optimal. First we note that at any time \( t \), the expected total satisfaction of the agent over the period \([t, T]\) depends on his current wealth \( W(t) \), his average past consumption \( z(t^-) \). It also depends on the current value of the risky assets prices \( S_n(t), n = 1, 2, \ldots, N \), since these prices summarize all the information about the future investment opportunities.

Let the risky securities have price \( S(\tau) \) at time \( \tau \). Assume that at time \( \tau \), which may be a random time, the agent has the opportunity to exchange his wealth and average past consumption for a level of satisfaction \( J^*(W(\tau), z(\tau^-), S(\tau), \tau) \). In other words, there is a "utility equivalent" function \( J^*(W, z, S, t) \) defined for all levels of wealth, average past consumption, risky asset prices, and for all times in \([0, T]\) that provides the agent at time \( t \) with lump-sum satisfaction for the period \([\tau, T]\). Therefore the agent has the option of choosing any budget feasible consumption plan till any time \( \tau \), and then exchanging his wealth and average past consumption at that time for a lump-sum payoff in utility units. We will show that if \( J^*(W, z, S, t) \) satisfies certain conditions, then exchanging \( W(0) \) and \( z(0^-) \) at time zero for \( J^*(W(0), z(0^-), S(0), 0) \) is better for the agent than adopting any feasible consumption plan till any random time \( \tau \) and then receiving a lump-sum utility at that time.

Consider the situation in which the agent decides to terminate at any random time \( \tau \) and receive a lump sum utility of \( J^*(W(\tau), z(\tau^-), S(\tau), \tau) \). In such a case, the total satisfaction of the agent will be given by:

\[ E\left[ \int_{0}^{\tau} u(z(s), s) ds + J^*(W(\tau), z(\tau^-), S(\tau), \tau) \right]. \quad (13) \]
To ensure that the value in (13) is the corresponding value for all possible plans, we impose the following boundary conditions on $J^*$. First, note that the agent might decide not to terminate a given consumption plan at all. At time $T$, he would receive $V(W(T^+))$ and his total satisfaction will given by (9). We require that at $T$, $J^*(W(T^+), z(T), S(T), T) = V(W(T^+))$, and thus the value in (13) is the same utility for a consumption plan followed completely till time $T$.

Second, note that the agent might follow a consumption plan that exhausts all his wealth strictly before time $T$, and then "live on" the satisfaction driven from past consumption. To ensure that the value given by (13) is the corresponding utility from such a plan, we impose the boundary condition that

$$J^*(0, z, S, t) = \int_t^T u(ze^{-\beta(s-t)}, s) \, ds + V(0).$$

After we prove that receiving $J^*$ is better for the agent than any other policy, we will provide sufficient conditions for an actual consumption and investment policy to have an associated value function equal to $J^*$. It then follows that such a policy will indeed be optimal. First, we need a definition and a technical lemma.

**Definition 2 (Smooth Manifolds)** Let $\phi_l(W, z, S, t) : \mathbb{R}^{N+2} \times [0, T] \to \mathbb{R}, l = 1, 2, \ldots, L$, where $L \leq N + 3$, be continuously differentiable functions. A smooth manifold $\mathcal{M}$ in $\mathbb{R}^{N+2} \times [0, T]$ is given by:

$$\mathcal{M} = \{(W, z, S, t) \in \mathbb{R}^{N+2} \times [0, T] : \phi_l(W, z, S, t) = 0, \quad l = 1, 2, \ldots, L\}. \quad (14)$$

**Lemma 1 (Generalized Itô's Lemma)** Fix an investment policy $A$, and a consumption policy $C$. Let $S(t), W(t), z(t)$ follow the dynamics in (4), (7) and (8), respectively. Let $f(W, z, S, t) : \mathbb{R}^{N+2} \times [0, T] \to \mathbb{R}$ be once continuously differentiable on $\mathbb{R}^{N+2} \times [0, T]$ in all of its arguments and twice continuously differentiable on $\mathbb{R}^{N+2} \times [0, T]$ in its first $N + 2$ arguments, except possibly on a smooth manifold $\mathcal{M}$. Then, for all optional times $\varrho \leq T$, we have:

$$f(W(\varrho), z(\varrho^-), S(\varrho), \varrho) = f(W(0), z(0^-), S(0), 0) + \int_0^\varrho [\mathcal{D}A f(s) + f_s(s)] \, ds$$

$$+ \int_0^\varrho [f_w(s) W(s) A(s)^T I_{S-s}(s) \sigma(s) + f_s(s)^T \sigma(s)] dB(s)$$
\[ + \int_0^\tau (\beta f_z(s) - f_W(s)) \, dC(s) \]
\[ + \sum_i \Delta f(\tau_i) - \sum_i [f_W(\tau_i) \Delta W(\tau_i) + f_z(\tau_i) \Delta z(\tau_i)] \text{ a.s.} \]

where \( \tau_1, \tau_2, \ldots \), are the jump points of \( C \), where all the partial derivatives are evaluated at the points \((W(t), z(t^-), S(t), t)\), where \( D^A f \) is the differential generator of \( f \) associated with the investment policy \( A \) and is given by:

\[
D^A f = f_W[W r + W A^T I_{S-1}(\mu - r S)] - f_z \beta z + f_z^T (\mu - \rho) \\
+ \frac{1}{2} [f_{WW} W^2 (A^T I_{S-1} \sigma)(A^T I_{S-1} \sigma)^T + 2 f_{WS}^T (W \sigma^T I_{S-1} A) + tr(f_{SS} \sigma^T \sigma)] .
\]

and where

\[
\Delta f(\tau_i) = f(W(\tau_i^+), z(\tau_i), S(\tau_i), \tau_i) - f(W(\tau_i), z(\tau_i^-), S(\tau_i), \tau_i).
\]

**Proof.** This follows from Krylov (1980, theorem 2.10.1) and Dellacherie and Meyer (1982, VIII.27). □

**Theorem 1** Assume that there exists a differentiable function \( J^*(W, z, S, t): \mathbb{R}^{N+2} \times \mathbb{R} \rightarrow \mathbb{R} \), concave in \( W \) and \( z \), which is continuously differentiable over \( \mathbb{R}^{N+2} \times [0, T] \) in all of its arguments, and twice continuously differentiable over \( \mathbb{R}^{N+2} \times [0, T] \) in its first \( N+2 \) arguments, except possibly on a smooth manifold \( M \), that satisfies the following differential inequality:

\[
\max \{ \max_A [u(z,t) + D^A J^* + J^*_W], \beta J^*_z - J^*_W \} = 0 \quad \text{for all} \quad t \in [0, T],
\]

together with the boundary conditions:

\[
J^*(W, z, T) = V(W) \quad \text{and} \\
J^*(0, z, t) = \int_t^T u(ze^{-\beta(s-t)}, s) \, ds .
\]

Then

\[
\mathbb{E} \left[ \int_0^\theta u(z(s), s) \, ds + J^*(W(\theta), z(\theta^-), S(\theta), \theta) \right] \leq J^*(W(0), z(0^-), S(0), 0) ,
\]

104
for all \((C, A) \in \mathcal{A}(W(0), 0)\) and all optional times \(\varrho \leq T\), where \(z(\varrho^-)\) and \(W(\varrho)\) are the average past consumption and wealth, respectively, associated with \((C, A)\), and where the space \(\mathcal{A}(W(0), 0)\) is the space of all budget feasible consumption plans, \(C\), and investment policies, \(A\), such that the process defined by the stochastic integral

\[
\int_0^t [J^*_W(s)W(s)A(s)\sigma(s) + J^*_z(s)\sigma(s)] dB(s)
\]

is a super-martingale.

**Proof.** Assume that the agent adopts a consumption plan \(C\) and an investment policy \(A\) which are members of \(\mathcal{A}(W(0), 0)\). Assume that he terminates \((C, A)\) at the optional time \(\varrho\). Let the points of discontinuity of \(C(\omega, t)\) on the stochastic interval \([0, \varrho]\) be \(\tau_1(\omega), \tau_2(\omega), \ldots\). Recall that \(W\) has left continuous and that \(z\) has right continuous sample paths. For simplicity of notation, we will at times denote \(J^*\) and its partial derivatives with respect to \(W, z, S\) and \(t\), evaluated at \((W(t), z(t^-), S(t), t)\), by \(J^*(t), J^*_W(t), J^*_z(t), J^*_S(t), J^*_t(t)\), respectively.

Lemma 1 implies that:

\[
\int_0^\varrho u(z(s), s) ds + J^*(W(\varrho), z(\varrho^-), S(\varrho), \varrho) = \int_0^\varrho u(z(s), s) ds + J^*(W(0), z(0^-), S(0), 0) + \int_0^\varrho [D^A J^*_W(s) + J^*_z(s)] ds
\]

\[
+ \int_0^\varrho [J^*_W(s)W(s)A(s)\sigma(s) + J^*_z(s)\sigma(s)] dB(s) + \int_0^\varrho [\beta J^*_z(s) - J^*_W(s)] dC(S) + \sum_i [J^*(\tau_i^+) - J^*(\tau_i)] - \sum_i [J^*_W(\tau_i)\Delta W(\tau_i) + J^*_z(\tau_i)\Delta z(\tau_i)] \quad P\text{-a.s.}
\]

By the hypothesis that \(J^*\) is concave in its first two arguments, we know that

\[
J^*(\tau_i^+) - J^*(\tau_i) \leq J^*_W(\tau_i)\Delta W(\tau_i) + J^*_z(\tau_i)\Delta z(\tau_i)
\]

\[
= -J^*_W(\tau_i)\Delta C(\tau_i) + J^*_z(\tau_i)\beta \Delta C(\tau_i) \quad \forall \tau_i \quad P\text{-a.s.}
\]

Furthermore, by the assumption that the quantity in (20) is a super-martingale, we conclude that:

\[
\mathbb{E} \left[ \int_0^\varrho [J^*_W(s)W(s)A(s)\sigma(s) + J^*_z(s)\sigma(s)] dB(s) \right] \leq 0.
\]

105
Putting these two conditions together, we conclude that

\[
E \left[ \int_0^\varrho u(z(s), s) \, ds + J^*(W(\varrho), z(\varrho^-), S(\varrho), \varrho) \right] \\
\leq J^*(W(0), z(0^-), S(0), 0) + E \left[ \int_0^\varrho \left[ u(z(s), s) + D^A J^*(s) + J^*_s \right] \, ds \\
+ \int_{0^-}^\varrho [\beta J^*_z - J^*_w] \, dC(s) \right] \\
\leq J^*(W(0), z(0^-), S(0), 0),
\]

where the second inequality follows from the assumption that both the integrands are negative and the fact that \( C \) is increasing. Note that the second integral on the right-hand side of the first inequality includes the possible jumps and the singular component of \( C \). The fact that (22) holds for all possible plans in \( A(W_0, 0) \) and for all optional times \( \varrho \leq T \) completes our proof. \( \Box \)

Now, if we can show that there exists an actual policy \((C^*, A^*)\) such that at any time \( t \) and starting from any state \((W(t), z(t^-), S(t))\), the policy is budget feasible, that is \((C^*, A^*) \in A(W(t), t)\), and the value function associated with this policy \( U(z(t^-), t; (C^*, A^*, W^*(T))) \) is equal to \( J^*(W(t), z(t^-), S(t), t) \), then we conclude from theorem 1 that the agent will prefer \((C^*, A^*)\) to any other consumption policy. In other words, \((C^*, A^*)\) will be his optimal choice.

In the following theorem, we provide a characterization of a policy \((C^*, A^*)\) which guarantees that the associated value function is equal to the function \( J^*(W, z, S, t) \) defined in theorem 1.

**Theorem 2** Suppose that there exists a differentiable function \( J^*(W, z, S, t) : \mathbb{R}^{N+2} \times [0, T] \to \mathbb{R} \cup \{ -\infty \} \), concave in \( W \) and \( z \), which is continuously differentiable over \( \mathbb{R}^{N+2} \times [0, T] \) in all of its arguments, and twice continuously differentiable over \( \mathbb{R}^{N+2} \times [0, T] \) in its first \( N+2 \) arguments, except possibly on a smooth manifold \( \mathcal{M} \), that satisfies the following differential inequality:

\[
\max \left\{ \max_A [u(z, t) + D^A J^* + J^*_t], \beta J^*_z - J^*_w \right\} = 0,
\]

(23)

together with the boundary conditions:

\[
J^*(W, z, T) = V(W) \quad \text{and} \quad \]

106
\[ J^*(0, z, t) = \int_t^T u(z e^{-\beta(t-s)}, s) \, ds + V(0). \]

Assume furthermore that, starting at any optional time \( \tau \in [0, T] \), with \( W(\tau), z(\tau^-), S(\tau) \), there exists a policy \((C^*, A^*) \in \mathcal{A}(W(\tau), \tau)\), for which \( C^* \) is continuous on \( (\tau, T] \), with a possible jump \( \Delta C^*(\tau) \), and whose associated state variables are \( W^*(t), z^*(t^-) \), such that for all \( t \in (\tau, T] \):

\[
\begin{align*}
    u(z^*, t) + D^{A^*} J^* + J^*_t &= 0 \quad \text{a.s.,} \quad (24) \\
    \mathbb{E} \left[ \left( \int_0^T \left[ J^*_W(s)W^*(s)A^*(s)^\top I_{S^{-1}}(s)\sigma(s) + J^*_S(s)\sigma(s)^\top \sigma(s) \right] ds \right)^{\frac{1}{2}} \right] &< \infty, \quad (25) \\
    \int_t^{t+\epsilon} [J^*_W(W^*, z^*, S, t) - \beta J^*_t(W^*, z^*, S, t)] dC^*(t) &= 0 \quad \text{a.s., } \forall \epsilon > 0, \quad (26)
\end{align*}
\]

and

\[
J^*(W(\tau), z(\tau^-), S(\tau), \tau) - J^*(W(\tau) - \Delta C^*(\tau), z(\tau^-) + \beta \Delta C^*(\tau), S(\tau), \tau) = 0 \quad \text{a.s.} \quad (27)
\]

Then \( U(z(t^-), t; (C^*, A^*, W^*(T))) = J^*(W(t), z(t^-), S(t), t) \) for all \((W(t), z(t^-), S(t), t) \in \mathbb{R}^{N+2} \times [0, T]\). In particular, \( U(z(0^-), 0; (C^*, A^*, W^*(T))) = J^*(W(0), z(0^-), S(0), 0) \). It then follows from theorem 1 that \((C^*, A^*)\) attains the maximum of \( U(z(0^-), 0; (C, A, W(T)))\) over the set \( \mathcal{A}(W(0), 0) \) as defined in theorem 1.

**Proof.** Recall from the definition of \( U(z(t^-), t; (C^*, A^*, W^*(T))) \) that for all stopping times \( \varrho_1 \) and \( \varrho_2 \) such that \( \tau \leq \varrho_1 < \varrho_2 \leq T \), we have:

\[
U(z^*(\varrho^-_1), \varrho_1; (C^*, A^*, W^*(T))) = \mathbb{E} \left[ \int_{\varrho_1}^{\varrho_2} u(z^*(s), s) \, ds + U(z^*(\varrho^-_2), \varrho_2; (C^*, A^*, W^*(T))) \big| \mathcal{F}_{\varrho_1} \right].
\]

(28)

Now consider the following:

\[
\begin{align*}
\mathbb{E} \left[ \int_{\varrho_1}^{\varrho_2} u(z^*(s), s) \, ds + J^*(W^*(\varrho_2), z^*(\varrho^-_2), S(\varrho_2), \varrho_2) \big| \mathcal{F}_{\varrho_1} \right] &= \mathbb{E} \left[ \int_{\varrho_1}^{\varrho_2} u(z^*(s), s) \, ds + J^*(W^*(\varrho^-_1), z^*(\varrho_1), S(\varrho_1), \varrho_1) + \int_{\varrho_1}^{\varrho_2} [D^{A^*} J^* + J^*_s] \, ds \\
&\quad + \int_{\varrho_1}^{\varrho_2} [J^*_W(s)W^*(s)A^*(s)^\top I_{S^{-1}}(s)\sigma(s) + J^*_S(s)\sigma(s)^\top \sigma(s)] dB(s) \\
&\quad + \int_{\varrho_1}^{\varrho_2} [\beta J^*_s - J^*_W] dC^*(s) \big| \mathcal{F}_{\varrho_1} \right]
\end{align*}
\]

107
\[
= J^*(W^*(\rho_1^+), z^*(\rho_1), S(\rho_1), \rho) + \mathbb{E} \left[ \int_{\rho_1}^{\rho_2} \left[ u(z^*(s), s) + D\lambda^* J^* + J^*_\rho \right] ds \right.
+ \left. \int_{\rho_1}^{\rho_2} \left[ \beta J_z^*(W^*, z^*, S, t) - J_W^*(W^*, z^*, S, t) \right] dC^*(s) \bigg| \mathcal{F}_{\rho_1} \right] \]

\[
= J^*(W^*(\rho_1^+), z^*(\rho_1), S(\rho_1), \rho_1) \]

\[
= J^*(W^*(\rho_1), z^*(\rho_1^-), S(\rho_1), \rho_1), \quad (29)
\]

where the first equality follows from lemma 1 and the assumption that \(C^*\) is continuous on \((\tau, T]\). The second equality follows from the facts: 1) that \(J^*(W^*(\rho_1^+), z^*(\rho_1), S(\rho_1), \rho_1)\) is \(\mathcal{F}_{\rho_1}\) measurable, 2) that the stochastic integral is a martingale. The fact that the stochastic integral is a martingale follows from (25). The third equality follows from equations (24) and (26), whereas the last equality follows from the continuity of \(C^*\) on \((\tau, T]\), and from (27) if \(\rho_1 = \tau\).

From equations (28) and (29), we conclude that for all stopping times \(\rho_1\) and \(\rho_2\) such that \(\tau \leq \rho_1 < \rho_2 \leq T\),

\[
U(z^*(\rho_1^-), \rho_1; (C^*, A^*, W^*(T))) - \mathbb{E}[U(z^*(\rho_2), \rho_2; (C^*, A^*, W^*(T)))|\mathcal{F}_{\rho_1}] = J^*(W^*(\rho_1), z^*(\rho_1^-), S(\rho_1), \rho_1) - \mathbb{E}[J^*(W^*(\rho_2), z^*(\rho_2), S(\rho_2), \rho_2)|\mathcal{F}_{\rho_1}].
\]

or

\[
\mathbb{E}[U(z^*(\rho), \rho; (C^*, A^*, W^*(T))) - J^*(W^*(\rho), z^*(\rho), S(\rho), \rho)|\mathcal{F}_{\tau}] = m,
\]

for all stopping times \(\rho\), and where \(m\) is a constant independent of \(\rho\). Now let \(\hat{T} = \min\{t \in [\tau, T]: W^*(t) = 0,\}\), where we have used the convention that if the minimum does not exist, we set the minimum to be \(T\). Note from the boundary condition on \(J^*\) that: \(J^*(W^*(\hat{T}), z^*(\hat{T}^-), S(\hat{T}), \hat{T}) = U(z^*(\hat{T}^-), \hat{T}; (C^*, A^*, W^*(T)))\) for all \(S(\hat{T})\). Thus we conclude that \(m = 0\), and the proof is now complete. \(\blacksquare\)

4 Optimal Consumption Policy

In this section we provide a closed form solution for a particular example of the optimal consumption problem formulated in section 2, in which the horizon is infinite. In this analysis, we use a felicity function \(u(z, t) = \frac{1}{\alpha}e^{-\theta t}z^{\alpha}\), where \(0 < \alpha < 1\), and where the
discount factor $\delta$ expresses the agent's impatience. In other words, the agent seeks to maximize

$$E\left[ \int_0^\infty e^{-\delta t} \frac{Z(t)^\alpha}{\alpha} \, dt \right],$$

where $Z$ is as defined in (8).

In addition, we assume that the prices of the risky securities are given by:

$$S(t) + \int_0^t \rho(s) \, ds = S(0) + \int_0^t I_S(s) \mu \, ds + \int_0^t I_S(s) \sigma \, dB(s), \quad (30)$$

where $I_S(t)$ is a diagonal $N \times N$ matrix whose diagonal elements are $S_n(t), n = 1, \ldots, N$, where $\mu$ is an $N \times 1$ vector of constants, and $\sigma$ is an $N \times M$ matrix of constants. The instantaneous interest rate $r$ is assumed to be a constant.

Note that since the asset price parameters and the interest rate are constant, the value function $J^*$ depends only on wealth $W$ and the average past consumption $z$. The maximum attainable utility starting in any state $(W, z)$ does not depend on the prices of assets, since the prices of assets do not contain any information about the investment opportunities in the future.

For a given investment policy $A$, which prescribes the ratio of wealth invested in each risky asset, the differential generator of any smooth function $J(W, z): \mathbb{R}_+ \rightarrow \mathbb{R} \cup \{-\infty\}$, see lemma 1, associated with $A$ is given by:

$$\mathcal{D}^A J = J_W [Wr + WA^T(\mu - r\mathbb{1})] - J_z \beta z + \frac{1}{2} J_{WW}[W^2 A^T(\sigma \sigma^T)A],$$

where $\mathbb{1}$ is an $N$-dimensional vector of ones. For simplicity of the notation, we will define:

$$\Gamma \equiv [\sigma \sigma^T]^{-1}[\mu - r\mathbb{1}] \quad \text{and} \quad \gamma \equiv [\mu - r\mathbb{1}]^T[\sigma \sigma^T]^{-1}[\mu - r\mathbb{1}]. \quad (31)$$

We will show that the optimal solution to the agent's problem takes the form of a ratio barrier policy. The optimal investment decision is to invest a constant fraction of wealth in the risky assets at all times. The optimal consumption is the amount required to keep the ratio of wealth $W$ to average past consumption $z$ less than $^6$ a critical number $k^*$ in

$^6$Recall that we use weak relations and hence $x$ less than $y$ means that $x$ is strictly less than or equal to $y.$

109
almost all states of nature. If the agent starts at time zero with wealth $W(0)$ and initial consumption experience $z(0^-)$ such that $\frac{W(0)}{z(0^-)}$ is strictly less than $k^*$, then the optimal consumption policy is to consume nothing and wait while $W$ increases on average and $z$ declines until the (random) time $\tau$ when the ratio $\frac{W(\tau)}{z(\tau)}$ is equal to $k^*$. Starting from that moment, the agent consumes only when $\frac{W}{z} = k^*$ the amount required to keep $\frac{W}{z} \leq k^*$ forever in all states of nature.

On the other hand, if the agent starts at time zero with $\frac{W(0)}{z(0^-)} > k^*$, then the optimal consumption policy is to take a "gulp" of consumption, reducing $W$ and increasing $z$, to bring the ratio $\frac{W}{z}$ immediately to $k^*$. Following this gulp, the optimal amount of consumption is that required to keep $\frac{W}{z} \leq k^*$ forever, and consumption occurs only when $\frac{W}{z} = k^*$. As we shall see, the erratic behavior of the price of the risky assets implies that the optimal consumption pattern will have "singular" sample paths. In almost all states of nature, the agent consumes a non trivial amount at infinitely many points of time. However, the times when consumption occurs are a set of Lebesgue measure zero.

We present the logic behind the optimal solution is two steps. First, we analyze general policies of the $k$-ratio barrier forms. For any $k > 0$, the corresponding $k$-ratio barrier policy is the policy of investing a constant fraction (to be given) of wealth in the risky assets together with a consumption policy of keeping the ratio $\frac{W}{z}$ less than or equal to $k$ forever in all states of nature. Consumption occurs only when $\frac{W}{z} = k$. This policy is followed after an initial "gulp" of consumption if the initial conditions are such that $\frac{W(0)}{z(0^-)} > k$, or after a (random) period of no consumption if $\frac{W(0)}{z(0^-)} < k$. We will show that the value function associated with any $k$-ratio barrier policy, denoted $J^k(W, z)$, satisfies

$$\max_A \left\{ \frac{z^\alpha}{\alpha} + D^A J^k - \delta J^k \right\} = 0 \quad \text{if} \quad W \leq k z \quad \text{and}$$

$$J^k_W - \beta J^k_z = 0 \quad \text{if} \quad W \geq k z .$$

Second, we will show that there is a unique value $k^* > 0$ such that the associated value function, $J^*(W, z)$, satisfies the differential inequality

$$\max_A \left\{ \max\left[ \frac{z^\alpha}{\alpha} + D^A J^* - \delta J^*, \ \beta J^*_z - J^*_W \right] \right\} = 0,$$

(33)

together with the boundary conditions that

$$J^*(0, z) = \frac{1}{\alpha} \int_0^\infty e^{-\delta t} (z e^{-\beta t})^\alpha \, dt \quad \text{and}$$

(34)

110
\[
\lim_{T \to \infty} E \left[ e^{-\delta T} J^*(W(T), z(T)) \right] = 0, \tag{35}
\]
as \(T\) ranges over nonstochastic times. It then follows from theorem 2, that the ratio barrier policy associated with \(k^*\) is the optimal solution for the agent’s problem.

To simplify the discussion, we will first provide a construction of the value function for any \(k\)--barrier policy assuming that the agent has no portfolio decision to make. In other words, we will assume that the agent can invest his wealth in one asset whose value \(S\) follows the following dynamics:

\[
dS = rS \, dt + \sigma S \, dB(t). \tag{36}
\]

After we construct the value function under these circumstances, we will analyze the case when the agent has to make a portfolio choice.

The object of control under a \(k\)--ratio policy is the ratio \(W/z\) and the purpose of control is to keep this ratio less than \(k\). Suppose that the agent starts from a state \((W, z)\) such that \(W/z < k\), and suppose that he invests all his wealth in the risky asset and consumes nothing until the stopping time \(\tau\) when \(W(\tau)/z(\tau) = k\). Also suppose that from \(\tau\) on, the agent consumes the minimum amount required to keep the ratio \(W(t)/z(t) \leq k\) for all \(t > \tau\). Let us compute the utility \(J^k(W, z)\) obtained from such a policy.

From the definition of the utility, we can write, for \((W, z)\) such that \(W/z < k\):

\[
J^k(W, z) = E \left[ \int_0^\tau e^{-\delta s} \frac{z^\alpha}{\alpha} e^{-\alpha \beta s} \, ds + e^{-\delta \tau} J^k(W(\tau), ze^{-\beta \tau}) \right] \\
= E \left[ \frac{z^\alpha}{\alpha(\alpha \beta + \delta)} [1 - e^{-(\alpha \beta + \delta) \tau}] \right] + E \left[ e^{-\delta \tau} J^k(W(\tau), ze^{-\beta \tau}) \right].
\]

We conjecture that for a point on the boundary \((W(\tau), z(\tau))\), we have

\[
J^k(W(\tau), z(\tau)) = cz(\tau)^\alpha, \tag{37}
\]
for some constant \(c\). Hence, we can rewrite \(J^k(W, z)\) as:

\[
J^k(W, z) = \frac{z^\alpha}{\alpha(\alpha \beta + \delta)} + \left[ c - \frac{1}{\alpha(\alpha \beta + \delta)} \right] z^\alpha E \left[ e^{-(\alpha \beta + \delta) \tau} \right]. \tag{38}
\]

Now, using the dynamics of wealth and average past consumption under the assumption of no consumption as long as \(W/z \leq k\), and using a computation due to Harrison (1985),
we can evaluate the term \( E[e^{-(\alpha \beta + \delta) r}] \). From Harrison (1985, proposition 23, page 42), we conclude that

\[
E[e^{-(\alpha \beta + \delta) r}] = e^{-q^*(\log k - \log \frac{W}{kz})} = (\frac{W}{kz})^{q^*}
\]

where

\[
q^* = \frac{1}{\sigma^2} \left[ \left( r + \beta - \frac{\sigma^2}{2} \right)^2 + 2\sigma^2(\alpha \beta + \delta) \right]^{\frac{1}{2}} + \left( r + \beta - \frac{\sigma^2}{2} \right),
\]

where \( q^* \) satisfies the quadratic equation:

\[
\left( r + \beta - \frac{\sigma^2}{2} \right)q^* + \frac{\sigma^2}{2}q^{*2} = \alpha \beta + \delta.
\]

Substituting from (39) into (38), we get that, when \( \frac{W}{z} < k \),

\[
J^k(W, z) = \frac{z^\alpha}{\alpha(\alpha \beta + \delta)} + \left[ c - \frac{1}{\alpha(\alpha \beta + \delta)} \right]z^\alpha \left( \frac{1}{kz} \right)^{q^*}.
\]

Now, consider the situation when \( \frac{W}{z} > k \). The \( k \)-barrier policy prescribes a jump of size \( \Delta = \frac{W-kz}{1+\beta k} \) to a point on the boundary and the corresponding utility is

\[
J^k(W, z) = e^{\left( r + \beta W \right)^\alpha (1 + \beta k)^\alpha}.
\]

Finally, we choose the value of the constant \( c \) such that \( J^k_W = \beta J^k_z \) when \( \frac{W}{z} = k \). From this condition, we get that

\[
c = \frac{\beta k}{(\alpha \beta + \delta)[q^*(1 + \beta k) - \alpha \beta k]} + \frac{1}{\alpha(\alpha \beta + \delta)}.
\]

To summarize, we conjecture that when there are no investment decisions to be made, the value function of the \( k \)-barrier policy is given by

\[
J^k(W, z) = \begin{cases} 
\frac{z^\alpha}{\alpha(\delta + \alpha \beta)} + z^\alpha \left[ \frac{W}{kz} \right]^{q^*} & \text{if } \frac{W}{z} \leq k \\
\left( \frac{z + \beta W}{1 + \beta k} \right)^\alpha & \text{if } \frac{W}{z} \geq k
\end{cases}
\]

where \( c \) is given in (44) and \( q^* \) is given in (40).

Now we turn to the original problem in which the agent has to choose, in addition to consumption, the proportions of wealth \( A^* \) to be invested in different assets. Suppose
that the agent follows also a $k$-ratio policy and that the corresponding value function has the following form when $\frac{W}{z} < k$:

$$J^k(W, z) = f_1(z) + f_2(z) W \alpha^*, \quad (45)$$

for some functions $f_1$ and $f_2$ and for a constant $\alpha^*$ to be determined. If this form were correct, then the optimal portfolio allocation $A^*$ will be given by

$$A^* = \frac{1}{1 - \alpha^*} \Gamma. \quad (46)$$

If the agent invests his wealth according to this rule, and assuming that he follows a $k$-ratio policy, then starting from a state $(W, z)$ such that $\frac{W}{z} < k$, the logarithm of the ratio $\frac{W}{z}$ will follow a Brownian motion with the a drift $\mu^*$ and a standard deviation $\sigma^*$ given by:

$$\mu^* = r + \beta + \frac{\gamma}{1 - \alpha^*} - \frac{\gamma}{2(1 - \alpha^*)^2}, \quad (47)$$

$$\sigma^* = \frac{\sqrt{\gamma}}{1 - \alpha^*}. \quad (48)$$

To determine the value of $\alpha^*$, we basically search for a fixed point. We should find the value of $\alpha^*$ with the property that the associated drift and variance terms of log $\frac{W}{z}$ produce a value function $J^k$ for the $k$-ratio policy such that the optimal investment policy computed from $J^k$ is $A^* \equiv \frac{\Gamma}{1 - \alpha^*}$. From Harrison (1985, equation 5, page 40), we conclude that $\alpha^*$ satisfies:

$$\mu^* \alpha^* + \frac{\sigma^* \alpha^2}{2} = \alpha \beta + \delta, \quad (49)$$

or

$$\alpha^* \left( r + \beta + \frac{\gamma}{2(1 - \alpha^*)} \right) = \alpha \beta + \delta. \quad (50)$$

We also need to make enough assumptions on the parameters of the problem to ensure strict concavity of the value function, and to ensure that $\alpha^*$ is a real number. These are recorded as follows:

**Assumption 2** The parameters of the problem satisfy

$$\delta > \alpha r \Rightarrow \frac{\alpha \gamma}{2(1 - \alpha)} \quad \text{and} \quad (1 - \alpha) \beta > \delta - r.$$

113
Given these restrictions, we can easily establish the following result.

**Lemma 2** Consider the quadratic equation

\[ x \left[ (r + \beta) + \frac{\gamma}{2(1 - x)} \right] = \alpha \beta + \delta, \]

and let the roots be \( x_1 \) and \( x_2 \), where \( x_2 > x_1 \). Under assumption 2, the roots are real, positive, \( x_2 > 1 \), and \( \alpha < x_1 < 1 \).

**Proof.** The reader can easily verify these properties. \( \square \)

**Remark 1** From now on, we will use \( \alpha^* \) to denote \( x_1 \), the smaller root of the quadratic equation in lemma 2.

Our conjecture takes its final shape when we combine the analysis of the value of \( \alpha^* \) with the previous discussion of the value function in the case of no investment decisions. In summary, we conjecture that the value function associated with any \( k \)-ratio policy is given by:

\[
J^k(W, z) = \begin{cases} 
\frac{z^{\alpha^*}}{c^{(x+\beta \gamma)^\alpha}} \left[ c - \frac{1}{\alpha (\alpha \beta + \delta)} \right] & \text{if } \frac{W}{z} \leq k \\
\frac{z^{\alpha^*}}{c^{(x+\beta \gamma)^\alpha}} & \text{if } \frac{W}{z} \geq k 
\end{cases}
\]

where \( \alpha^* \) is given in lemma 2 and

\[
c = \frac{\beta k}{(\alpha \beta + \delta) \left[ \alpha^* (1 + \beta k) - \alpha \beta k \right]} + \frac{1}{\alpha (\alpha \beta + \delta)}. \tag{51}
\]

Now, fix \( k > 0 \) and consider the ratio barrier policy associated with \( k \). The following proposition confirms our conjecture and proves that the supposed value function is in fact the utility that the agent attains when he follows a \( k \)-ratio policy.

**Proposition 1** Consider the function \( J^k(W, z) : \mathbb{R}_{+}^2 \to \mathbb{R} \cup \{-\infty\} \) given by

\[
J^k(W, z) = \begin{cases} 
\frac{z^{\alpha^*}}{c^{(x+\beta \gamma)^\alpha}} \left[ c - \frac{1}{\alpha (\alpha \beta + \delta)} \right] & \text{if } \frac{W}{z} \leq k \\
\frac{z^{\alpha^*}}{c^{(x+\beta \gamma)^\alpha}} & \text{if } \frac{W}{z} \geq k 
\end{cases}
\]

where \( \alpha^* \) is given in lemma 2 and \( c \) is given in (51).
The function \( J^k \) is continuous, strictly concave, has continuous first derivatives, has second derivatives and satisfies, together with \( A^* = \frac{1}{1-\sigma} \Gamma \):

\[
\frac{z^\alpha}{\alpha} + D^\alpha J^k - \delta J^k = 0 \quad \text{if} \quad W \leq kz \quad \text{and} \quad (52)
\]

\[J^k_W - \beta J^k_z = 0 \quad \text{if} \quad W \geq kz, \quad (53)\]

and the boundary conditions

\[J(0, z) = \frac{1}{\alpha(\alpha \beta + \delta)} z^\alpha \quad \text{and} \quad (54)\]

\[\lim_{T \to \infty} E[e^{-\delta T} J^k(W(T), z(T))] = 0, \quad (55)\]

as \( T \) ranges over nonstochastic times.

**Proof.** All the assertions can be verified by direct computations using the definition of \( \alpha^* \). Strict concavity of \( J^k \) in \( W \) follows from the fact that \( \alpha^* < 1 \), and in \( z \) follows from the fact that \( \alpha < \alpha^* \). For the last boundary condition note that on any feasible path \( W(t) \leq kz(t) \) for all \( t > 0 \), and hence \( J^k(W(T), z(T)) \) is less than \( z^\alpha(T) \) multiplied by a constant. But

\[
E[z(T)] = E[z(0^-)e^{-\beta T} + \beta \int_0^T e^{-\beta(T-s)} dC(s)]
\]

\[
< z(0^-)e^{-\beta T} + E[\beta C(T)].
\]

But from the budget constraint in (7), we can easily see that \( E[C(T)] \geq E[\hat{W}(T)] \) where \( \hat{W}(T) \) is the wealth realized when following portfolio rule \( A^* \) with no consumption withdrawal before \( T \). Using the dynamics of \( \hat{W} \), and the properties of \( \alpha^* \), the reader can easily verify that

\[
\lim_{T \to \infty} E[e^{-\delta T} J(W(T), z(T))] \leq \lim_{T \to \infty} E[e^{-\delta T} z^\alpha(T) c] \leq \lim_{T \to \infty} E[c e^{-\delta T} [z(0^-)e^{-\beta T} + \beta \hat{W}(T)]^\alpha]
\]

where \( c \) is given by equation (51). But

\[
[z(0)e^{-\beta T} + \beta \hat{W}(T)]^\alpha \to [\beta \hat{W}(T)]^\alpha \quad \text{uniformly in} \quad \omega \quad \text{as} \quad T \uparrow \infty
\]

hence

\[
\lim_{T \to \infty} E[e^{-\delta T} (z(0)e^{-\beta T} + \beta \hat{W})^\alpha] = \lim_{T \to \infty} E[e^{-\delta T} \beta^\alpha \hat{W}^\alpha(T)] \quad \text{and thus}
\]

\[
\lim_{T \to \infty} E[e^{-\delta T} J(W(T), z(T))] \leq \lim_{T \to \infty} E[e^{-\delta T} \beta^\alpha \hat{W}^\alpha(T)].
\]
But
\[
\mathbb{E}[\hat{W}(T)]^\alpha = W(0) e^{\alpha \left[ r + \frac{\gamma(1 - 2\alpha^*)}{2(1 - \alpha^*)^2} + \frac{\alpha^* \gamma^2}{2(1 - \alpha^*)^2} \right] T} < W(0) e^{\alpha^* \left[ r + \frac{\gamma(1 - 2\alpha^*)}{2(1 - \alpha^*)^2} + \frac{\alpha^* \gamma^2}{2(1 - \alpha^*)^2} \right] T} \leq W(0) e^{\alpha^* (r + \beta) + \frac{\alpha^* \gamma^2}{2(1 - \alpha^*)^2} T - \alpha^* \beta T} \leq W(0) e^{(\alpha \beta + \delta) T - \alpha^* \beta T},
\]
where the first inequality follows from the property that \( \alpha < \alpha^* \), and the last inequality from the definition of \( \alpha^* \). Hence \( \mathbb{E}[e^{-\delta T} \hat{W}^\alpha(T)] \leq W(0) e^{\beta (\alpha - \alpha^*) T} \). Noting that \( \alpha - \alpha^* < 0 \), the required result follows.

Now, we show that the function \( J^k \) developed in proposition 1 is indeed the utility that the agent gets from following a \( k \)-ratio barrier policy.

**Proposition 2** The function \( J^k(W, z) : \mathbb{R}_+^2 \rightarrow \mathbb{R} \cup \{-\infty\} \) given in proposition 1 is the value of following a \( k \)-ratio barrier policy.

**Proof.** Fix any initial point \((W, z)\). We need to show that for the \( k \)-ratio barrier policy
\[
J^k(W, z) = \mathbb{E}\left[ \int_0^\infty e^{-\delta t} \frac{z^\alpha(t)}{\alpha} dt \right] .
\]
Remember that if \( W > kz \), a "gulp" of size \( \Delta \) is consumed and \( J^k(W, z) \equiv J^k(W - \Delta, z + \beta \Delta) \). Also, recall that after the possible initial jump, the consumption policy satisfies
\[
\int_0^T (J^k_W - \beta J^k_z) dC = 0 \quad a.s.
\]
for any deterministic time \( T \). Now, from lemma 1, we can write
\[
\mathbb{E}\left[ \int_0^T e^{-\delta t} \frac{z^\alpha}{\alpha} dt + e^{-\delta T} J^k(W(T), z(T)) \right] = J^k(W, z) + \Delta J^k + \mathbb{E}\left[ \int_0^T e^{-\delta t} \left( \frac{z^\alpha}{\alpha} + \mathcal{D} \alpha^* J^k - \delta J^k \right) dt \right] + \mathbb{E}\left[ \int_0^T e^{-\delta t} (J^k_W - \beta J^k_z) dC(t) \right] = J^k(W, z),
\]
where \( \Delta J^k \) is the jump in \( J^k \) at time zero. The second equality follows from the observations that, by construction, \( \Delta J^k = 0 \); that \( J^k \) satisfies, together with \( A^* \), equation (52);
and that for the barrier policy consumption occurs only when $J^*_W = \beta J^*_z$. The assertion follows by taking the limit of both sides in the above equation as $T \uparrow \infty$, and using condition (55).

This proposition shows that any ratio barrier policy satisfies many of the sufficient conditions for optimality. All policies, except one, fail to satisfy the differential inequality (33). The one barrier policy that satisfies inequality (33) is the optimal solution, which is recorded in the following proposition.

**Proposition 3** Let
\[ k^* = \frac{1 - \alpha^*}{\beta(\alpha^* - \alpha)}, \]
and note that the $k^* > 0$ by lemma 2. The ratio barrier policy associated with $k^*$ is the optimal solution for the agent's problem.

**Proof.** Let the value function associated with the $k^*$-ratio barrier policy be $J^*$. In view of proposition 1, we only need to show that $A^* = \frac{1}{1 - \alpha^*} \Gamma$ and $J^*$ satisfy:
\[
J^*_W - \beta J^*_z \leq 0 \quad \text{if} \quad W \leq k^*z \quad \text{and} \quad \frac{z^\alpha}{\alpha} + \mathcal{D} A^* J^* - \delta J^* \leq 0 \quad \text{if} \quad W \geq k^*z.
\]

We start with the first inequality. Using the definition of $J^*$ provided in proposition 1, we can write $J^*_W - \beta J^*_z$ in the region $W \leq k^*z$ as:
\[
J^*_W - \beta J^*_z = z^{\alpha - 1} \mathcal{H} g\left(\frac{k^*_2}{W}\right) \quad \text{where} \quad g(y) = \beta(\alpha^* - \alpha)y^{\alpha^*} + \frac{\alpha^*}{k^*} y^{(\alpha^* - 1)} - \frac{\beta}{\alpha \beta + \delta},
\]
where $\mathcal{H}$ is a constant. Note that $g(1) = 0$, and that since $\alpha^* > 0$, $g(y) \uparrow \infty$ as $y \uparrow \infty$.

Computing the derivative of $g$, we get
\[
\frac{dg}{dy} = \alpha^* \beta(\alpha^* - \alpha)y^{\alpha^* - 1} + \frac{\alpha^*(\alpha^* - 1)}{k^*} y^{\alpha^* - 2}.
\]
Substituting $k^*$ in the expression for $\frac{dg}{dy}$, we conclude that $\frac{dg}{dy} = 0$ for $y = 1$, and that $\frac{dg}{dy} > 0$ for $y > 1$. It then follows that $J^*_W - \beta J^*_z \geq 0$, for all points $(W, z)$ such that $W \leq k^*z$.

117
Now, we verify the second inequality. Consider any point \( \mathbf{x} \equiv (W, z) \) such that \( W \geq k^* z \). Referring to Figure 1, let \( \mathbf{a} \) be the point on the intersection of the line \( W = k^* z \) and the straight line passing through the point \( \mathbf{z} \) with slope \( \frac{dW}{dz} = -\frac{1}{\beta} \). In otherwords, \( \mathbf{a} \) is the point on the boundary \( W = k^* z \), to which one would jump if one starts at \( \mathbf{z} \). By construction, \( J^*(\mathbf{x}) = J^*(\mathbf{a}) \).

Let

\[
f(W, z) = \frac{z^\alpha}{\alpha} + J^*_r W - \frac{J^*_{W^2}}{2J^*_W} \gamma - J^*_z \beta z,
\]

and note that \( f(\mathbf{a}) - \delta J^*(\mathbf{a}) = 0 \). Applying the fundamental theorem of calculus along the straight line connecting \( \mathbf{a} \) and \( \mathbf{x} \), we obtain

\[
f(\mathbf{x}) - f(\mathbf{a}) = \int_{\mathbf{a}}^{\mathbf{x}} f_W \, dW + f_z \, dz.
\]

Noting that along the line connecting \( \mathbf{a} \) and \( \mathbf{x} \), we have \( dW = -\beta dW \), and that \( dW > 0 \) in the direction from \( \mathbf{a} \) to \( \mathbf{x} \), it then follows that \( f_W - \beta f_z \leq 0 \) is sufficient to conclude that

\[
\frac{z^\alpha}{\alpha} + J^*_r W - \frac{J^*_{W^2}}{2J^*_W} \gamma - J^*_z \beta z - \delta J^* \leq 0 \quad \text{if} \quad \frac{W}{z} \geq k^*.
\]

Computing \( f_W - \beta f_z \) in the region \( W \geq k^* z \), and using the properties of \( \alpha^* \), the reader can easily verify that \( f_W - \beta f_z \leq 0 \) for \( k^* = \frac{1 - \alpha^*}{\beta(\alpha^* - \alpha)} \).

The following proposition records the optimal consumption and investment policy.

**Proposition 4** Let

\[
A^*(t) = \frac{1}{1 - \alpha^*} \Gamma \quad \text{for all} \ t,
\]

and let the budget feasible consumption process \( C^* \), which has continuous sample paths almost surely, be given by:

\[
C^*(t) = \Delta C^*(0) + \int_0^t \frac{W^*(s)}{1 + \beta k^*} \, dl(s) \quad P - a.s.
\]

where

\[
\Delta C^*(0) = \max \left\{ 0, \frac{W(0) - k^* z(0^-)}{1 + \beta k^*} \right\},
\]

\[
l(t) = \sup_{0 \leq s \leq t} \left[ \log \frac{W(s)}{\hat{z}(s)} - \log k^* \right]^+ \quad P - a.s.,
\]

\[118\]
\[ \hat{W}(t) = (W_0 - \Delta C^*(0))e^{([r+\frac{\beta^2}{2}\frac{1}{1-\alpha^2}]t+[\frac{1}{1-\alpha^2}\Gamma^\top\sigma]B(t))} P - a.s., \quad (59) \]
\[ \hat{z}(t) = (z_0 + \beta \Delta C^*(0))e^{-\beta t} P - a.s., \quad (60) \]

and where \( W^*(s) \) and \( z^*(s) \) are the state variables associated with \( C^* \). The strategy \( A^* \) and \( C^* \) defined above are the optimal solution for the agent’s problem.

To prove this proposition, all we need to show is that the above strategy is the ratio barrier strategy associated with \( k^* \). The \( k^* \) ratio barrier policy calls for a jump at \( t = 0 \), if \( \frac{W(0)}{z(0)} > k^* \). The size of the jump \( \Delta \) is calculated to achieve the condition \( \frac{W(0)-\Delta}{z(0)+\beta \Delta} = k^* \), from which we compute \( \Delta \) as in (57).

Let us proceed now to compute the consumption process after \( t = 0 \). The \( k^* \)-ratio barrier policy is equivalent to the condition that

\[ \frac{W^*(t)}{z^*(t)} \leq k^* \quad \forall t, P - a.s. \quad \text{or} \]
\[ \log \frac{W^*(t)}{z^*(t)} \leq \log k^* \quad \forall t, P - a.s. \quad (62) \]

Now consider the “unregulated” process \( \hat{W}, \hat{z} \), which gives the wealth of the agent and his past average consumption under the assumption that no consumption takes place after the initial jump at \( t = 0 \). The values of \( \hat{W}(t) \) and \( \hat{z}(t) \) are given in (59) and (60). The ratio \( \frac{W(t)}{z(t)} \) will fail to satisfy condition (61) for some periods and for some sample paths, and the idea of the solution is to “regulate” this ratio using the consumption process \( C^* \) to ensure that at the optimal solution condition (61) is satisfied.

To achieve this regulation, define the process

\[ l(t) = \sup_{0 \leq s \leq t} \left[ \log \frac{\hat{W}(s)}{\hat{z}(s)} - \log k^* \right]^+ P - a.s. \]

and let the “regulated” process \( \log \frac{W^*}{z^*} \) be given by:

\[ \log \frac{W^*(t)}{z^*(t)} = \log \frac{\hat{W}(t)}{\hat{z}(t)} - l(t) P - a.s. \quad (63) \]

For each state \( \omega \), the sample path \( l(\omega, .) \) has the following properties:

* \( l(\omega, .) \) is increasing and continuous with \( l(\omega, 0) = 0 \), for almost all \( \omega \).
\[ \log \frac{W^*(t)}{z^*(t)} \leq \log k^* \text{ for all } t \geq 0, \text{ P-a.s.} \]

\[ l(\omega, .) \text{ increases only when } \log \frac{W^*(\omega, t)}{z^*(\omega, t)} = \log k^*. \]

This "regulated" process (log \( \frac{W^*(t)}{z^*(t)} \)) is a candidate for the logarithm of the ratio of the state variables associated with the optimal consumption plan, since from the above construction, we can easily conclude that condition (61) is satisfied. The question now becomes whether there exists a feasible consumption process \( C^* \) that could enforce the relationship in (63). Expanding both sides of (63) using Itô’s lemma, we get

\[
\log \frac{W^*(t)}{z^*(t)} = \log \frac{W(0) - \Delta C^*(0)}{z(0^-)} + \beta \Delta C^*(0) + \int_t^t \left[ r + \beta + \frac{\gamma(1-2\alpha^*)}{2(1-\alpha^*)^2} \right] ds \\
+ \int_0^t \frac{1}{(1-\alpha^*)} \Gamma^T \sigma dB(s) - \int_0^t \left[ \frac{1}{W^*(s)} + \frac{\beta}{z^*(s)} \right] dC^*(s) \quad P - a.s.
\]

Thus, we conclude that the consumption process given by:

\[
\int_t^t \left[ \frac{1}{W^*(s)} + \frac{\beta}{z^*(s)} \right] dC^*(s) = l(t) \quad P - a.s. \tag{64}
\]

satisfies the condition in (63). Note from the above equation that \( C^* \) increases only when \( l \) increases for all sample paths. Therefore, \( C^* \) increases only when \( \frac{W^*(t)}{z^*(t)} = k^* \) and hence condition (26) of theorem 2 is satisfied.

Now let the process \( y \) be given by \( y(t) = \frac{W^*(t)z^*(t)}{\beta W^*(t) + z^*(t)} \). From Itô’s lemma, we get:

\[
y(t)l(t) = y(0)l(0) + \int_0^t y(s) dl(s) + \int_0^t l(s) dy(s) \quad P - a.s.
\]

Noting that \( l(0) = 0, P - a.s., \) and that \( dl(s) = \frac{dC^*(s)}{\nu(s)} \), we find that

\[
y(t)l(t) = C^*(t) + \int_0^t l(s) dy(s) \quad P - a.s.
\]

Integrating the second term in the right-hand side by parts, we have that

\[
C^*(t) = \int_0^t \frac{W^*(s)z^*(s)}{\beta W^*(s) + z^*(s)} \, dl(s) \quad P - a.s.
\]

120
But from the properties of \( l \), we know that \( l \) increases only when \( \frac{W^*(s)}{Z^*(s)} = k^* \). We thus conclude that

\[
C^*(t) = \int_0^t \frac{W^*(s)}{1 + \beta k^*} dl(s) \quad P - a.s.
\]

The solution we constructed has the following features. After the initial (possible) gulp, the agent consumes the minimum possible amount required to keep the marginal ratio of wealth equal to (\( \beta \) times) the marginal value of \( z \). Consumption takes place only when these marginal values are equalized. If the marginal value of wealth drops below (\( \beta \) times) the marginal value of \( z \), the agent stops consuming. These rules result in a consumption process with singular sample paths. Consumption occurs at uncountably infinite number of times, but the set of all times when consumption occurs has Lebesgue measure zero, for almost all sample paths. This result occurs because of the unbounded variation property of the Brownian sample paths.

In short, we have produced a solution composed of a possible initial jump followed by an continuous consumption process with singular sample paths that satisfies all the sufficient conditions stated in theorem 2. Hence, we conclude that this solution is the \textit{unique} solution to the agent's problem starting from any values of \( W(0), z(0^-) \).

The reader can easily see from the above construction that the state space is divided into the following regions:

\[
\mathcal{O} = \{(W, z) \in \mathbb{R}_+ \times \mathbb{R}_+: W < k^*z\},
\]

\[
\mathcal{B} = \{(W, z) \in \mathbb{R}_+ \times \mathbb{R}_+: W = k^*z\},
\]

\[
\mathcal{R} = \{(W, z) \in \mathbb{R}_+ \times \mathbb{R}_+: W > k^*z\}.
\]

The object of control is the location of the point \( (\hat{W}(t), \hat{z}(t)) \) which is subject to random shocks because of the randomness of the return on the risky assets. The purpose of control is to keep this point inside the region \( \mathcal{O} \) be consuming only when it reaches the boundary \( \mathcal{B} \). The sample path properties of the Brownian motion lead to peculiar behavior of the point \( (W(t), z(t)) \) when it reaches the boundary \( \mathcal{B} \). Assume, for now, that the agent does not consume at all. He just invests his wealth optimally. Let this "unregulated" process of wealth and average past consumption be \( \hat{W} \) and \( \hat{z} \), respectively. When the point \( (\hat{W}(t), \hat{z}(t)) \) reaches the boundary \( \mathcal{B} \) it will hit it uncountably infinite number of times and will move back and forth between the regions \( \mathcal{O} \) and \( \mathcal{R} \) uncountably.
infinite number of times before it leaves the neighborhood of the boundary. The total time that \((\hat{W}(t), \hat{z}(t))\) actually spends on the boundary is of Lebesgue measure zero. This phenomenon happens because of the unbounded variation nature of the sample paths of Brownian motion.

In order to control this "unregulated" process and produce the optimal process \((W^*(t), z^*(t))\) which lies always in \(\mathcal{O} \cup \mathcal{B}\), the agent needs to apply the control, consumption, when the process \((\hat{W}(t), \hat{z}(t))\) is on the boundary. Furthermore, the agent consumes the least possible amount required to keep \((\hat{W}(t), \hat{z}(t))\) on the boundary. The result is a process of cumulative consumption with a nontrivial increasing sample paths. However, the points of increase of each sample path are of Lebesgue measure zero. In other words, the sample paths of the optimal consumption process are singular functions: nontrivial increasing functions whose derivatives are zero for almost all \(t\). This property is due to the very erratic nature of the sample paths of Brownian motion.

5 Concluding Remarks

In this essay, we have provided sufficient conditions for a consumption and portfolio policy to be optimal for a class of time-nonseparable preferences that consider consumptions at nearby dates to be almost perfect substitutes. We demonstrated our general theory by explicitly solving in closed form the optimal consumption and portfolio policy for a particular felicity function when the prices of the risky assets follow a geometric Brownian motion process.

The optimal consumption policy in our solution consists of a possible initial "gulp" of consumption followed by a consumption and investment behavior in which consumption occurs only periodically. The resulting process of cumulative consumption has singular sample paths. That is: it has sample paths which are increasing nontrivial functions whose derivatives are zero almost everywhere. Furthermore, the points of increase of each sample path have Lebesgue measure zero. Problems of singular control have been studied by many authors in contexts different from ours. We refer the reader to Harrison (1985) and the references therein.
6 References

1. K. Arrow and M. Kurz, Public Investment, the Rate of Return, and Optimal Fiscal Policy, Johns Hopkins Press, Baltimore, 1970.


The State Space showing the Boundary B. For all times after \( t=0 \), the optimal policy restricts the state to the admissible region \( O \).

**Figure 1**
Essay IV

On Risk Aversion in Dynamic Models

Abstract

We survey the literature on dynamic choice theory and discuss measures of risk aversion in a dynamic framework. We summarize the difficulties of measuring attitudes towards risk when utility is derived from many commodities. We present Kreps and Porteus (1978) development of preferences over temporal lotteries and Seldien's (1978) ordinal certainty equivalent approach. We survey recent applications to asset pricing, notably Epstein and Zin (1989) contributions, and we provide a critique of their interpretations. Finally, we point out directions for further research.
1 Introduction and Summary

Preferences for risky prospects about which uncertainty is gradually resolved are critical determinants of the choices of individuals and the outcomes of dynamic economic activity. For economic theorists, three important features of preference representation seem essential: 1) support by well formulated axioms about individual choice, 2) analytical tractability in simple models with testable implications, and 3) empirical Support.

Recent developments in the literature of dynamic choice in general, and consumption/investment choice in particular, emphasized some of the above mentioned aspects at the expense of others. For example, the "standard" state separable/time additive model of preferences with homogeneous one period utility function is tractable and provides useful insights about individual choice; see Merton (1971), and market equilibrium; see Cox, Ingersoll and Ross (1985). However, the assumption of time separability is very strong since it implies that consumptions at two adjacent dates are perfect non-substitutes. Furthermore, recent empirical studies; see Heaton (1990) for a summary, do not lend support for this specification.

On the other hand, many investigations which have their origins in decision theory; see Kreps and Porteus (1978) and Selden (1978), develop preferences with very strong axiomatic foundations. Nevertheless, such preferences are not widely used in a parametric form in the context of asset prices and market equilibrium, with the notable exception of Epstein and Zin (1989).

The purpose of this essay is to provide a synthesis of the existing literature on dynamic choice with a special focus on the notion of risk aversion in a dynamic framework and with the intention of discovering areas of useful further research. We organize our synthesis around the famous "substitution axiom" used by von Neumann and Morgenstern to develop expected utility preferences. This essay can be thought of as providing answers to the following questions:

- What is the relation between risk aversion and preferences on sure bundles in a world with many commodities in which agents satisfy the independence axiom?

- When does the independence axiom fail? and how do we model such circumstances?
• What is the content of the substitution axiom? and should our models of primitive preferences move in the direction of relaxing or imposing the substitution axiom?

We provide the answer to the first question in section 2. We show that within the class of preferences over gambles whose prizes are vectors of commodities which satisfy the substitution axiom, one can not meaningfully separate preferences for certain consumption from attitudes towards risk by studying primitive preferences over consumption lotteries. One can use the indirect utility function to study induced preferences on wealth gambles provided that all consumption decisions are made after uncertainty about wealth is resolved. We demonstrate this by some examples. We also review the results of Stiglitz (1969) who shows that assumptions about risk attitudes for gambles on timeless wealth imply restrictions on the indifference curves for sure consumption.

To answer the second question, we present the issue raised by Dreze and Modigliani (1972) about the timing of choice in section 3. Dreze and Modigliani (1972) pointed out that when decisions about consumption have to be made before all uncertainty about wealth is resolved, the induced preferences over wealth lotteries fail to satisfy the independence axiom. This is an example of what can be called the phenomenon of "temporal risk". We then review Kreps and Poleus (1978) "induced" preferences in which the timing of resolution of uncertainty is of prime concern.

In section 4, we review the analysis by Selden (1978) for preferences over "certain ¥ uncertain" consumption pairs, whose motivation is to develop parametric representation of preferences in which preferences for risk and for intertemporal substitution are parameterized independently. We also present the necessary axiom, the "coherence axiom", which needs to be imposed on the structure proposed by Selden to obtain expected utility representation.

In section 5, we review the work of Epstein and Zin (1989) of recursive utility function. We provide a critique of the view that their structure is an extension of Kreps and Poleus preferences, and we argue against their interpretations of attitudes towards the timing of resolution of uncertainty. In section 6, we summarize the different approaches to modeling intertemporal preferences and we argue that: 1) preferences for consumption be modeled as primitive rather than induced, and hence the timing of resolution of uncertainty need not be a determinant of preferences, and 2) relaxing time additivity and maintaining
state separability is more intuitive than relaxing state separability and maintaining time additivity. Finally, we point out directions for further research.

2 Risk Aversion with Many Commodities

Suppose that we have two consumers who are faced with the choice between lotteries on a vector of consumption goods. Can we, in general, describe their attitudes towards risk by observing their choices of lotteries? The answer is, in general, no. This can be seen from the following example which is adapted from Kilhstrom and Mirman (1974).

Let \( u \) and \( v \) be two distinct utility functions representing two different preference orderings on the set \( \{(x, y) \in \mathbb{R}_+^2\} \), where \( x \) is consumption today and \( y \) is consumption tomorrow. Let \( z \equiv (x, y) \) and \( z' \equiv (x', y') \) be two distinct points in \( \mathbb{R}_+^2 \), with the property that \( u(z) > u(z') \) and \( v(z') > v(z) \), as shown in figure 1. Now suppose that consumer 1, with utility function \( u \), and consumer 2, with utility function \( v \), are both faced with the choice of receiving \( z \) with certainty or a gamble on \( z \) and \( z' \). It is evident that consumer 1 will choose \( z \) with certainty, while consumer 2 will choose the gamble. Consumer 1 is not more "risk averse" than consumer 2; he just has different ordinal preferences on consumption today and consumption tomorrow.

One can try to express risk aversion by studying induced preferences on income rather than primitive preferences on consumption. We can easily show that if preferences on uncertain consumption bundles satisfy the von Neuman-Morgenstern (NM) axioms, and if all consumption choices are made after uncertainty about income is resolved, then induced preferences on income also satisfy NM axioms and can be expressed as expectation of the indirect utility function. One can then apply the Arrow-Pratt measure of risk aversion to the indirect utility function for purposes of comparative statics.

Stiglitz (1969) shows that the indirect utility function is restricted by the preferences on the underlying consumption and that assumptions about the indirect utility function, such as risk neutrality or constant relative risk aversion, are also assumptions about primitive preferences. We reproduce one of his theorems in the following.

Suppose that one assumes that an agent has a choice from \( n \) consumption goods with prices given by the vector \( p \). The indirect utility function of wealth, \( V(W; p) \), is the
maximum attainable utility given wealth \( W \) and prices \( p \). Suppose that \( V(W; p) \) exhibits constant relative risk aversion, then \( V(W; p) \) can be shown to be

\[
V(W; p) := g(p)W^{\gamma(p)+1} + h(p), \quad \gamma(p) \neq -1,
\]

(1)

where \(-\gamma(p)\) is the degree of relative risk aversion for a given level of prices, and where \( g \) and \( h \) are functions of prices. The income -consumption curves corresponding to \( V \) are given by:

\[
x^i = -\frac{g_i W}{W(\gamma + 1)} - \frac{\gamma_i}{(\gamma + 1)} W \log W - \frac{h_i W^{-\gamma}}{W(\gamma + 1)},
\]

(2)

where \( x^i \) is the demand for good \( i \), and where we used subscripts to denote partial derivatives. Stiglitz then shows that if (2) is to hold for all levels of wealth the indifference map over certain consumption must be homothetic.

We can summarize this literature in the following observation:

**Observation 1** Within the framework of NM preferences over a space of many commodities, one can not meaningfully separate preferences for certain consumption from attitudes towards risk by studying primitive preferences over consumption lotteries. One can use the indirect utility function to study induced preferences on wealth gambles provided that all consumption decisions are made after uncertainty about wealth is resolved. One can then use Arrow-Pratt measures of risk aversion on the indirect utility function. These measures, however, will depend on the environment, for example on prices of commodities. Furthermore, attitudes towards risk as expressed by the indirect utility function are restricted by preferences on the underlying consumption goods.

### 2.1 Intertemporal Consumption

In a multi-period world, a commodity which may be consumed at \( n \) different times can be modeled as \( n \) different commodities. The preference ordering on the space of sure intertemporal consumption, coupled with NM assumptions on gambles of consumption bundles, produces definite forms of risk aversion on timeless gambles of wealth. The following examples illustrate this point. Consider an investor who can allocate his wealth between a riskless asset with instantaneous riskless return \( r \) and a risky asset whose
price $S$ follows a logarithmic Brownian motion with drift parameter $\mu$ and local variance parameter $\sigma$. In other words,

$$dS = \mu S + \sigma S d\omega,$$

where $\omega$ is a standard Brownian motion. In addition, the investor plans to consume during his lifetime.

**Example 1 (Merton (1971))** Merton considers time-separable preferences for consumption, given by:

$$U^m = E\left[\int_0^\infty e^{-\delta t} \frac{c(t)^\alpha}{\alpha} dt\right],$$

where $c(t)$ is the consumption rate at time $t$, and where $\alpha < 1$. In addition, one needs to impose the condition that

$$\delta > \alpha r + \frac{\alpha}{2(1 - \alpha)} \left(\frac{\mu - r}{\sigma}\right)^2.$$

Merton (1971) shows that the indirect utility function for wealth, given $(r, \mu, \sigma)$, is

$$V^m(W) = \left(\frac{1}{1 - \alpha}\right)^{\alpha - 1} \left[\frac{\alpha}{2(1 - \alpha)} \left(\frac{\mu - r}{\sigma}\right)^2\right]^{\alpha - 1} \frac{W^\alpha}{\alpha}.$$

Therefore, the Arrow-Pratt measure of relative risk aversion to timeless gambles on wealth is $R^m_t = 1 - \alpha$. In addition, Merton shows that the proportion of wealth invested in the risky asset is $\frac{1}{R^m_t} \left(\frac{\mu - r}{\sigma^2}\right)$. The higher the degree of relative risk aversion for timeless gambles on wealth, the lower the proportion of wealth invested in the risky asset at each point in time.

**Example 2 (Hindy and Huang (1990))** Hindy and Huang (1990) study the behavior of an investor who faces the same investment opportunities but who has different preferences on sure intertemporal consumption. The agent derives satisfaction not only from current consumption but also from past consumption. In particular, let

$$z(t) = z(0^-) e^{-\beta t} + \beta \int_0^t e^{-\beta(t-s)} dC(s),$$

where $C(t)$ is the total consumption till time $t$, $z(0^-)$ is initial consumption experience, and where $\beta$ captures the rate of decay of the effect of past consumption on current satisfaction. The agent maximizes

$$U^{h-h} = E\left[\int_0^\infty e^{-\delta t} \frac{z(t)^\alpha}{\alpha} dt\right],$$

131
where $\alpha < 1$. In addition to condition (4), one needs to assume that $\beta > \frac{\delta - r}{1-\alpha}$. Hindy and Huang (1990) show that the indirect utility function for wealth for such an agent is given as follows. There are two positive constants $\alpha^*$ and $k^*$ given by:

$$\alpha^* = \frac{1}{2} \left[ 1 + \frac{\gamma}{2(r + \beta)} + \frac{\alpha \beta + \delta}{r + \beta} \right] - \frac{1}{2} \left[ 1 + \frac{\gamma}{2(r + \beta)} + \frac{\alpha \beta + \delta}{r + \beta} \right]^2 - 4 \left( \frac{\alpha \beta + \delta}{r + \beta} \right)^{\frac{1}{2}}$$

and

$$k^* = \frac{1 - \alpha^*}{\beta(\alpha^* - \alpha)},$$

where $\gamma = (\frac{\mu - \tau}{\sigma})^2$. Given the assumptions on the parameters of the problem, one can easily check that $\alpha < \alpha^* < 1$.

The indirect utility function, which depends on the state variables $(W,z)$, is given by:

$$V^{h-h}(W,z) = \begin{cases} 
Ax^\alpha + z^\alpha \left[ \frac{W}{z} \right]^\alpha B & \text{if } \frac{W}{z} \leq k^* \\
(z + \beta W)^\alpha C & \text{if } \frac{W}{z} \geq k^*
\end{cases}$$

where $A, B$ and $C$ are constants. Computing the Arrow-Pratt measure of relative risk aversion to timeless gambles on wealth, $R^{h-h}_r$, we get:

$$R^{h-h}_r(W,z) = \begin{cases} 
1 - \alpha^* & \text{if } \frac{W}{z} \leq k^* \\
\frac{1 - \alpha^*}{1 + \alpha \frac{W}{z}} & \text{if } \frac{W}{z} \geq k^*
\end{cases}.$$

Hindy and Huang (1990) also show that the proportion of wealth invested in the risky asset is $\frac{1}{1-\alpha^*}(\frac{W}{z})^{\alpha^*}$.

Note that for two agents with preferences specified as in (6), and with the same parameters $(\alpha, \beta, \delta)$, and the same wealth $W$, the measure of risk aversion depends on the ratio $\frac{W}{z}$. The lower this ratio is, the lower the measure of relative risk aversion. In addition, an agent with preferences as in (6) will exhibit smaller relative risk aversion for timeless gambles on wealth than an agent with time additive preferences with the same wealth and with the same curvature, $\alpha$, of the instantaneous utility function.

As we discussed before, there are two crucial assumptions required to use the indirect utility functions as a measure of risk:

- Primitive preferences on consumption gambles satisfy NM axioms.

132
• All decisions about consumption allocation are made after uncertainty about wealth is resolved.

These two assumptions imply that induced preferences on wealth satisfy NM axioms and that they can be represented by the indirect utility function. In addition, preferences for intertemporal consumption will affect the measure of risk. There are two distinct bodies of literature that analyze preferences when either of these two assumptions is violated.

When decisions about consumption have to be made before all uncertainty about wealth is resolved, the induced preferences fail to satisfy the independence (substitution) axiom of NM. This was pointed out by Dreze and Modigliani (1972). As a result, researchers have studied induced preferences which violate the independence axiom and which take account of the timing of the resolution of uncertainty. The seminal work in this area is by Kreps and Porteus (1978). Another line of research, pioneered by Selden (1978), studies primitive preferences on consumption which violate the NM independence axiom in an attempt to separate attitudes towards risk from attitudes towards intertemporal substitution. We review, briefly, these results in the following sections.

3 Kreps and Porteus Preferences

3.1 Temporal risk and induced preferences

Dreze and Modigliani (1972) pointed out that when decisions about consumption have to be made before all uncertainty about wealth is resolved, the induced preferences over wealth lotteries fail to satisfy the independence axiom of NM. This is an example of what can be called the phenomenon of "temporal risk". The fundamental characteristic of situations in which choice has to be made before uncertainty is completely resolved is that the decision maker is likely to have to make some other decisions in the mean time. Dreze and Modigliani (1972) coined the term "temporal risk" for such a situation, and they called the interim actions the "auxiliary decisions". The alternative case, when uncertainty is immediately resolved or when the agent faces no auxiliary decisions is referred to as that of "timeless risks".

To see how individuals choose in a situation of temporal risk, consider an individual facing a set of alternative temporal prospects \( \{x_t\} \) with respective distribution functions
\{F_i(.)\}, and who must make some auxiliary choice \(\alpha\) out of a set \(\mathcal{A}\). Assume that the agent seeks to maximize the expectation of a NM utility function \(\phi(x, \alpha)\). In such a case, the individual ranks each temporal prospect on the basis of how much expected utility it will produce given the corresponding optimal auxiliary decision. In other words, the agent ranks lotteries on the basis of the value:

\[
V(F_i) \equiv \int \phi(x, \alpha^*(F_i)) \ dF_i(x) \quad \text{where} \\
\alpha^*(F_i) \equiv \arg\max_{\alpha \in \mathcal{A}} \int \phi(x, \alpha) \ dF_i(x).
\]

The function \(V(F_i)\) is the individual’s "induced preference" functional over temporal distributions and the ranking it generates represents the agent’s induced preferences. In the case of timeless risks, the choice of \(\alpha\) could be postponed till after all the relevant uncertainty is resolved. In this case, the agent ranks prospects on the basis of the value

\[
W(F_i) \equiv \int \phi(x, \alpha^*(x)) \ dF_i(x) \quad \text{where} \\
\alpha^*(x) \equiv \arg\max_{\alpha \in \mathcal{A}} \phi(x, \alpha).
\]

It is clear that the functionals \(V\) and \(W\) are, in general, not ordinally equivalent.

The difference between \(V\) and \(W\) is exactly the difference between induced preferences in cases of temporal and timeless risks. Note that \(W\) is a linear functional of \(F(.)\) and hence induced preferences satisfy the independence or substitution axiom, and can be represented as the expectation of NM utility function \(\nu(x) \equiv \phi(x, \alpha^*(x))\). On the other hand, \(V\) is a nonlinear functional of the distribution functions. This is why the induced preferences fail to satisfy the independence axiom. As a consequence, there does not exist, in general, NM utility function capable of summarizing the ranking generated by \(V\).

3.2 The approach

The above section demonstrates that temporal risks lead to induced preferences which violate the independence axiom. The implications of this observation for modelling decision making in a dynamic model are quite important. If temporal, as opposed to timeless, prospects are the rule rather than the exception, then the systematic violation of linearity
outlined before implies that using expected utility models might lead to systematically incorrect results.

One way to avoid this problem is to model explicitly the auxiliary decisions that have to be made before all uncertainty is resolved. One can then analyze the jointly optimal risky choices and auxiliary decisions. Auxiliary decisions, however, are multi-dimensional, differ across individuals, and are often of little interest to the researchers. Kreps and Porteus (1978, p.83) conclude that “the obvious difficulty with this approach is that such complete models may become overburdened with detail and analytically intractable”. Since it is theoretically incorrect to use expected utility approach and usually impractical to leave all auxiliary decisions in a model, they conclude that they “are forced to develop preference structures which can be used to model induced preferences directly”.

We review very briefly Kreps and Porteus (KP) preferences here. The preferences are defined on temporal lotteries.

3.3 Temporal lotteries

We restrict our attention to a two-period model. An agent receives payoffs at each of two times \( t = 0, 1 \). The payoff at time \( t \) is denoted \( z_t \) and is drawn from a compact subset of a complete separable metric space \( Z_t \). The uncertainty about \( z_0 \) must resolve at time 0, whereas the uncertainty about \( z_1 \) may resolve at either \( t = 0 \) or \( t = 1 \). To formalize this, let \( D_1 \) be the space of all Borel probability measures on \( Z_1 \), with generic element \( d_1 \). Let \( D_0 \) be the set of all probability measures on \( Z_0 \times D_1 \), with generic element \( d_0 \). Elements of \( D_0 \) are called temporal lotteries. Temporal lotteries are lotteries in which the timing of resolution of uncertainty is encoded.

3.4 Kreps and Porteus axioms

Kreps and Porteus assume the following axioms on a preference relation \( \succ \) on the space \( D_0 \) of temporal lotteries:

1. Continuity in the weak convergence topology.

2. Independence at \( t = 0 \). If \( d_0 \succ d_0'' \) and \( \alpha \in (0, 1) \), then \( \alpha d_0 + (1 - \alpha)d_0'' \succ \alpha d_0' + (1 - \alpha)d_0'' \), for all \( d_0'' \in D_0 \).
3. Independence at $t = 1$. Let $(z_0, d_1)$ be the degenerate measure in $D_0$ with prize $(z_0, d_1)$. If $(z_0, d_1) \succ (z_0, d'_1)$ and $\alpha \in (0, 1)$, then $(z_0, \alpha d_1 + (1 - \alpha)d''_1) \succ (z_0, \alpha d'_1 + (1 - \alpha)d''_1)$, for all $d''_1 \in D_1$.

Kreps and Porteus show that the above axioms lead to the following result.

**Proposition 1 (Kreps and Porteus (1978))** The above axioms are necessary and sufficient for there to exist continuous functions $U_1: Z_0 \times Z_1 \to \mathbb{R}$ and $u_0: Z_0 \times \mathbb{R} \to \mathbb{R}$ such that $u_0$ is strictly increasing in its second argument and if $U_0: Z_0 \times D_1 \to \mathbb{R}$ is defined by

$$U_0(z_0, d_1) = u_0(z_0, \mathbb{E}[U_1(z_0, z_1); d_1]),$$

then $d_0 \succ d'_0$ if and only if $\mathbb{E}[U_0(z_0, d_1); d_0] > \mathbb{E}[U_0(z_0, d_1); d'_0]$.

Note that there are two "utility functions" in the above representation of preferences: $U_1$ for uncertainty which resolves at time one and $U_0$ for uncertainty which resolves at time zero. The function $u_0$ acts to convert "time one utility levels" to corresponding "time zero utility levels". This conversion is not just a "rescaling" of utility, it rather embodies attitudes towards the timing of resolution of uncertainty.

### 3.5 Preference for early or late resolution

Kreps and Porteus (1978) show that if a representation of preferences over temporal lotteries is possible with $u_0(z_0, r)$ affine in $r$, then preferences can be represented by standard NM utility on the vector $(z_0, z_1)$. This is the case if and only if the above axioms hold, and

$$(z_0, \alpha d_1 + (1 - \alpha)d'_1) \sim \alpha(z_0, d_1) + (1 - \alpha)(z_0, d'_1),$$

for all $z_0, d_1, d'_1$, and $\alpha$. In such a case, preferences are said to be resolution neutral. Other special cases are when preferences are resolution seeking:

$$(z_0, \alpha d_1 + (1 - \alpha)d'_1) \prec \alpha(z_0, d_1) + (1 - \alpha)(z_0, d'_1),$$

for all $z_0, d_1, d'_1$, and $\alpha$, and when preferences are resolution averse:

$$(z_0, \alpha d_1 + (1 - \alpha)d'_1) \succ \alpha(z_0, d_1) + (1 - \alpha)(z_0, d'_1),$$

136
for all $z_0, d_1, d_1', \text{ and } \alpha$. A necessary and sufficient condition for (12) is that $u_0$ is convex in its second argument. Necessary and sufficient for (13) is that $u_0$ is concave in its second argument.

4 Selden’s Ordinal Certainty Equivalent (OCE) Preferences

4.1 Structure

Selden (1978) proposed a representation of preferences over “certain \times uncertain” consumption pairs as an alternative to NM preferences on vectors of today and tomorrow’s consumption. Realizing that NM preferences impose strong restrictions on attitudes towards risk from preferences over sure consumption bundles, Selden’s motivation was to develop representations of preferences in which attitudes towards risk can be modelled independently from intertemporal preferences for sure consumption.

In Selden’s framework, an agent has, for each value of today’s sure consumption, conditional “risk” preferences for tomorrow’s uncertain consumption which satisfy the NM axioms. The choice between pairs of certain consumption today and uncertain consumption tomorrow can be decomposed into two steps. First, these pairs can be transformed into pairs of certain consumption today and a certainty equivalent consumption tomorrow using NM expected utility representation for tomorrow’s uncertain consumption. Second, the latter pair can be ordered using an ordinal time preference relation defined on certain consumption plans. In other words, Selden’s preferences are based on a conditional expected utility function for consumption tomorrow and a two-period ordinal index as in the following example.

Example 3 (Selden(1978)) Suppose that an agent confronts a risky consumption/savings problem. Let his “time” preferences be represented by

$$U(c_1,c_2) = (-c_1^{-\delta_1})/\delta_1 + (-c_2^{-\delta_1})/\delta_1, \quad -1 < \delta_1 < \infty,$$

and his (conditional) “risk” preferences by a period-two expected utility function with the constant relative risk aversion NM index

$$V(c_2) = -c_2^{-\delta_2}/\delta_2, \quad -1 < \delta_2 < \infty.$$
Then $\delta_1$ and $\delta_2$ are interpretable, respectively, as "time" and "risk" parameters. Using Selden preferences theory, $\delta_1$ and $\delta_2$ can be prescribed separately and their effects on optimal saving and consumption be studied. In contrast, using NM preferences on the vector of today-tomorrow consumption, $\delta_1$ must equal $\delta_2$, and hence the separate role of time and risk preferences can not be distinguished.

4.2 The Coherence Axiom

Rossman and Selden (1978) asked the following question: under what conditions do OCE preferences satisfy NM axioms? In other words: what is the content of NM assumptions on the space of today-tomorrow consumption pairs? The answer is contained in what they call the "coherence" axiom, which we describe in the following.

Recall that OCE preferences, denoted $\succsim$, are composed of two preference relations—a "time" preference relation, $\succsim^{\xi}$, on certain today-tomorrow consumption, and a family of conditional risk preferences $\{\succsim^{\xi} | x \in \mathbb{R}_+\}$ over uncertain consumption tomorrow, where $x$ is consumption today. Denote the implied indifference relations by $\sim, \sim^{\xi}$ and $\sim^{\xi}$, respectively. Rossman and Selden (1978) define a transfer mapping $\lambda: C[x] \rightarrow C[x']$, where $C[x] \equiv \{x\} \times \mathbb{R}_+$, characterized by the relation $c \sim^{\xi} \lambda c$, for each $c$ in the domain of $\lambda$, see Figure 2. The induced transfer mapping associates to each pair $(x, F)$, where $F$ is a probability distribution on tomorrow's consumption, the pair $(x', G)$ where $F(y) = G(y')$ if $\lambda(x, y) = (x', y')$.

Starting from a pair $(x, F)$, one obtains the corresponding pair $(x', G) = \lambda(x, F)$ by "sliding along the intertemporal indifference curves". For instance, suppose that $F$ has $n$ jump points $\{y_1, \cdots, y_n\}$. Then under the induced transfer, one constructs a c.d.f $G$ which:

- has $n$ jump points $\{y'_1, \cdots, y'_n\}$, where each $y'_i$ is obtained by finding $(x', y'_i)$ on the same indifference curve as $(x, y_i)$.

- has the same probability structure as $F$, or $F(y_i) = G(y'_i)$.

Conditional risk preferences are said to be coherent, if for two pairs of "certain× uncertain" consumption with the same value of today's consumption, $s$ and $s'$, $s \sim s'$ implies that $\lambda s \sim \lambda s'$. The coherence axiom simply states that the transfer map $\lambda$, 138
which is completely determined by "time" preferences, maps a conditional indifference set in $S[x]$ into a conditional indifference set in $S[x']$, where $S[x]$ is $\{x\}$ times the set of probability distributions on tomorrow's consumption. Coherence is thus a property linking time and conditional risk preferences. Rossman and Selden (1978) show that the coherence axiom is essentially what is required to transform OCE preferences into NM preferences on today-tomorrow consumption pairs.

It is important to understand further the implications of the coherence axiom. The axiom implies that conditional risk preferences over tomorrow's consumption are invariant under a transformation based on time preferences. In other words, starting with a pair $(x, F)$, and changing the allocations of consumption along the indifference curves while keeping the probability structure constant has no effect on preferences in the sense that $\lambda(x, F) \sim (x, F)$. This property has a strong intuitive appeal.

We can explain this point differently. Consider the case of the pairs $(x, F)$ and $(x', G)$ in Figure 3. Assume that the conditional preferences for tomorrow's consumption, corresponding to $x$ and $x'$, are given, respectively, by $V_x$ and $V_{x'}$. One can compute the conditional certainty equivalent value of consumption tomorrow as:

$$\hat{y} = V_x^{-1}(\pi V_x(y_1) + (1 - \pi)V_x(y_2)),$$

$$\hat{y}' = V_{x'}^{-1}(\pi V_{x'}(y_1) + (1 - \pi)V_{x'}(y_2)).$$

On the basis of conditional risk preference, $(x, \hat{y})$ is indifferent to $(x, F)$, and $(x', \hat{y}')$ to $(x', G)$. Refer now to Figure 4. It is easy to see that if $(x, y_i) \sim (x', y'_i), i = 1, 2$, then coherence requires that $(x, \hat{y})$ and $(x', \hat{y}')$ lie on the same time preference indifference curve. This also holds for all lotteries. In contrast, Selden's OCE preferences would, in general, allow an indifference curve passing through $(x, \hat{y})$ to lie above or below $(x', \hat{y}')$.

5 Applications in Asset Pricing

Expected utility preferences for multi-period consumption with an intertemporally additive and homogeneous one period (or instantaneous) utility function of the form

$$U = E\left[\int_0^T \frac{c(t)^\alpha}{\alpha} dt\right] \text{ where } \alpha < 1 \quad (14)$$
have been utilized to study consumption and asset demands. They have also been utilized in general equilibrium models to study asset prices. Analytical tractability is obtained at the cost of making very strong assumptions about preferences. It has been realized that an unsatisfactory feature of the specification (14) is that two distinct aspects of preferences, risk aversion and intertemporal substitution, are closely related. In fact, the coefficient of relative risk aversion for gambles on timeless wealth is the reciprocal of the elasticity of substitution. In addition, expected utility representative agent, optimizing models based on (14) did not perform well empirically; see Heaton (1990) for a brief discussion.

Epstein and Zin (1989), Weil (1990), and Farmer (1990) developed a class of preference specification in which the elasticity of substitution and risk aversion can be parameterized independently. Duffie and Epstein (1990) extend the specification to a continuous time environment. We review here the elements of the construction introduced by Epstein and Zin (1989), henceforth abbreviated as EZ, which include the following:

5.1 Consumption Space

The consumption space of EZ is a subset of temporal lotteries on consumption in a discrete time infinite horizon environment. EZ adopt the following notation: For any metric space $X$, denote by $B(X)$ the Borel $\sigma$-field and by $\mathcal{M}(X)$ the space of Borel probability measures on $X$ endowed with the weak convergence topology.

A temporal lottery $d$ can be pictured as an infinite probability tree in which each branch corresponds to a deterministic consumption stream $y \in \mathbb{R}_+^\infty$. Denote by $\mathcal{D}$ the space of such lotteries endowed with some metric. The lottery $d$ can be identified with a pair $(c_0, m)$, where $c_0 \geq 0$ denotes the nonstochastic period zero level of consumption and $m$ is a probability measure over the set of $t = 1$ nodes in the tree. In other words, $m$ can be thought of as an element of $\mathcal{M}(D)$. EZ define their consumption space as a subset of $\mathcal{D}$, denoted $\mathcal{D}^*$, in which the consumption streams in each branch of the tree are "bounded" in some appropriate sense. We refer the reader to Epstein and Zin (1989) for details.
5.2 Certainty Equivalent Functional and Aggregator

EZ define a certainty equivalent functional $\mu : \mathcal{M}(\mathbb{R}_+) \rightarrow \mathbb{R}_+$ which is consistent with first and second degree stochastic dominance. In addition, $\mu$ has the property that if a gamble yields the outcome $x$ with certainty, then the certainty equivalent of the gamble is $x$. EZ use the certainty equivalent functional together with an aggregator function $W$ to construct recursive utility $V$ in two steps. First, for any temporal lottery $(c_0, m)$, the random future utility is transformed into a certainty equivalent utility using $\mu$. In the second step, the future utility is combined with $c_0$ via the aggregator $W$.

Formally, given a utility function $V : D^* \rightarrow \mathbb{R}_+$, and $(c_0, m) \in D^*$, denote by $V[m]$ the probability measure for future utility implied by $V$ and $m$, in the sense that:

$$V[m](Q) = m \{ d \in D^* : V(d) \in Q \}, \quad Q \in \mathcal{B}(\mathbb{R}_+).$$

The utility function $V$ is called recursive if it satisfies the following equation on its domain:

$$V(c_0, m) = W(c_0, \mu(V[m])), \quad (16)$$

for some increasing aggregator function $W : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$, and some certainty equivalent functional $\mu$.

5.3 Recursive Utility

EZ restrict their attention to aggregators, $W$, of the form

$$W(c, z) = [c^\rho + \beta z^\rho]^{1\rho}, 0 \neq \rho < 1, 0 < \beta < 1.$$ 

(17)

In the case of deterministic consumption programs, $V$ is an intertemporal constant elasticity of substitution utility function with elasticity of substitution $\frac{1}{1-\rho}$. Using this class of aggregators, EZ provide sufficient conditions for the existence of a recursive utility function $V$ that satisfies (16). Furthermore, they show that such specification is tractable in the sense that dynamic programming can be used to characterize the optimal consumption and portfolio choice for an agent with preferences given by $V$. Finally, they derive asset pricing implications using a model of a representative, utility maximizing agent with preferences given by $V$. They show, in particular, that the systematic risk of
an asset is determined by covariance with both the return to the market portfolio and consumption growth. In more traditional models, like CAPM or Consumption based CAPM, only one of these factors is sufficient to characterize asset returns.

5.4 A Critique of EZ interpretations

The preference specification introduced by EZ in (16) and (17) has a recursive structure reminiscent of KP specification in (10). However, the construction of EZ is more in the spirit of the two step procedure recommended by Selden (1978) and recorded in section 4.1. In such a procedure, one first computes the certainty equivalent of tomorrow's utility using a certainty equivalent functional, and then combines it with today's consumption using an aggregator.

The procedure of computing a certainty equivalent is not always consistent with KP algorithm, in which tomorrow's expected utility, rather than a certainty equivalent, is combined with today's consumption via a scaling function $u_0$; see proposition 1. Therefore, when one uses the certainty equivalent approach, one can not use the properties of the aggregator $W$ to analyze preferences for the timing of resolution of uncertainty. In other words, EZ aggregator, $W$, is not equivalent to KP scaling function $u_0$.

To see this point more clearly, consider the preference specification introduced by EZ (1989) in section 3, and labeled "class II: Kreps/Porteus". The certainty equivalent functional is given by:

$$\mu(p) \equiv (E_p x^\alpha)_{\sigma}, \quad p \in \mathcal{M}(\mathbb{R}_+),$$

where $0 \neq \alpha < 1$. Recall that the aggregator $W$ is given by equation (17), and hence the recursive utility $V$ satisfies

$$V(c_0, m) = \left[ c_0^\alpha + \beta(E_m V^\alpha(.))_{\sigma}^{\frac{1}{\sigma}} \right]^\frac{1}{\alpha} .$$

Note that (19) does not correspond to KP representation in (10), since the scaling function, $u_0$, in KP algorithm should have $(EV)$ as its second argument, rather than $(EV^\alpha)$. EZ rewrite (19) in the form

$$U(c_0, m) = H(c_0, E_m U(.)) ,$$

142
where $U \equiv \frac{V^\alpha}{\alpha}$ is ordinally equivalent to $V$ and where

$$H(c, z) \equiv \left[ e^c + \beta(\alpha z)^{\frac{\rho}{\alpha}} \right]^{\frac{1}{\rho}} / \alpha.$$  

EZ then use this version of preferences to analyze, in section 4, attitudes towards risk and preferences for the timing of resolution of uncertainty, based on the results of KP as summarized in section 3.5.

The problem with this logic is that the transform $U \equiv \frac{V^\alpha}{\alpha}$ required to obtain KP representation is not an affine transform, and hence (20) does not produce the same preferences, in KP framework, as (19). This results because the scaling function $u_\alpha$, or in this context $H$, as defined by KP, is a cardinal index determined up to affine transforms. For a proof, see Kreps and Porteus (1978), or recall that their preferences are assumed to satisfy the independence axiom at each point in time.

The discussion in EZ (1989), Farmer (1990) and Weil (1990) about preferences for timing of resolution, although intuitively plausible, can not be justified by appealing to Kreps and Porteus results. In particular, using the representation (19), EZ argue that the certainty equivalent functional (18) is the embodiment of risk aversion, since it characterizes gambles on timeless wealth. Next, shifting to the representation (20), they argue that the ratio $\frac{\rho}{\alpha}$ is the determinant of the attitudes towards timing of resolution of uncertainty.

According to Epstein and Zin (1989, p.952), "the characterization of the KP class raises an issue which we suspect is relevant more generally and which calls for some attention. We have interpreted $\alpha$ as a risk aversion parameter. But with $\rho$ fixed, a reduction in $\alpha$ not only increases risk aversion but also may transform a preference for late resolution into a preference for early resolution. One is left wondering how to interpret the comparative statics effects of a change in $\alpha$. Similarly, a change in $\rho$ for given $\alpha$ affects both substitutability and attitudes towards timing. Thus the latter aspect of preference seems intertwined with both substitutability and risk aversion".

The above argument is flawed since it relies on the interpretation of the parameters $\alpha$ and $\rho$ in two different equations, (19) and (20), which do not represent the same preferences. To wit, the certainty equivalent approach of EZ requires an ordinal transform of the recursive utility to cast it into KP framework. Unfortunately, such transforms are not order preserving on the space of temporal lotteries.
I believe that the best interpretation of EZ (1989) preferences is to think of their recursive structure as a generalization of the functional form introduced by Selden (1978) to infinite horizon consumption programs. One can study the properties of such a specification, as they did. However, the axiomatic foundation for such preferences is not yet fully analyzed. Furthermore, the reliance on KP results to interpret the parameters of the recursive utility is not justified precisely because EZ rely on a non-order-preserving transform.

6 Future Research

6.1 Theoretical Work

A useful way of relating the different approaches to modelling preferences for consumption in a dynamic context is depicted in Figure 5. First, one should decide whether to model preferences over consumption streams as primitive or induced. As we explained earlier, the distinction between primitive and induced preferences hinges on whether “other” decisions have to be made before all uncertainty about consumption is resolved.

Kreps and Porteus (1979) have the view that preferences over intertemporal consumption should be modelled as induced preferences. They contend that “using expected utility maximization for consumption streams is suspect. Consumption, and more generally all economic activity, takes place within a larger context of social and political activities. Therefore preference for consumption alone is an induced preference. If the optimal social/political decision today changes with changes in next year’s consumption, then earlier resolution of uncertainty concerning next year’s consumption will be valuable, and this value should be reflected in models of preference for consumption”.

One could argue, on the other hand, that the sequence of decisions runs in the other direction. One’s choice of consumption follows rather than precedes the social and political choices. If one takes this view, then thinking of preferences for consumption as a primitive concept which does not affect other decisions is a reasonable route to take. Following this route, one needs to decide whether to model preferences in an expected utility framework. As Rossman and Selden (1978) show, and as we will elaborate, the hypothesis of expected utility representation implies a strong link between preferences
over sure consumption bundles and preferences over consumption gambles. This relation is the content of the coherence axiom discussed in section 4.2.

The behavioral content of the expected utility specification is the condition that if an agent is faced with a consumption lottery, and if the prizes in each state are replaced by other prizes which the agent thinks are equivalent under conditions of certainty, then the new lottery is equivalent to the original one. In other words, a state-by-state exchange of equivalent prizes produces equivalent consumption lotteries. This condition is intuitively appealing since an agent will obtain the prize in only one state after the uncertainty is resolved. It seems natural that changing the lottery by offering him an equivalent prize in that state will make him indifferent between the original and the new gamble.

The unsatisfactory aspects of the traditional specification of intertemporal preference with time and state separable utilities is, in my judgement, because of the assumption of time separability and not because of state separability. The dimensions of time and state are very different economically. Time has a natural ordering and there is a very well defined sense of distance between two time points. One also feels that consumptions at two adjacent time points should be treated as almost perfect substitutes. On the other hand, the states of nature have no natural economic order or notion of closeness. States of nature are all mutually exclusive, only one state will be eventually realized. In contrast, all moments of time will be experienced. It hence seems reasonable to pursue the analysis of preferences in the direction of relaxing time separability and maintaining state separability. This is the approach taken by Hindy and Huang (1989). In contrast, Epstein and Zin (1989) and Duffie and Epstein (1990) take the approach of relaxing state separability while maintaining time separability.

The choice of which route to follow, relaxing time versus state separability, is a matter of taste and intuition. At this stage it seems to me that the relaxing time separability is the more appealing approach. However, more work needs to be done. The class of preferences introduced by Hindy and Huang (1989) should be studied to produce asset pricing implications in an equilibrium framework. On the other hand, the axiomatic foundation of EZ and Duffie and Epstein (1990) preferences should be developed since as we argued in section 5.4, using Kreps and Porteus (1979) axioms is not justified. One should study Selden's (1978) preferences in a multi-period and in a continuous time
setting.

6.2 Empirical Work

One of the motivations of Selden's (1978) approach and Epstein and Zin (1989) later extension is to develop parameteric models of preferences to replace the standard specification in (3) in which the parameters of risk aversion can be separated from the parameters for intertemporal substitution. These models deliver a separation by directly introducing a new parameter through the certainty equivalent functional. However, this is achieved at the cost of relaxing state separability and maintaining time separability. On theoretical ground, it seems that the opposite approach of relaxing time separability while maintaining state separability is more plausible.

Recently, the parametric models of asset demand and consumption choice studied by Constantinides (1988), Heaton (1989), Hindy and Huang (1990), among others, include a new parameter; $\beta$ in example 2, and produce values of risk aversion that are not directly related to coefficients of intertemporal substitution. Such models enrich the standard specification in (3) and provide a new degree of freedom in modeling.

One important judge of the applicability of the time-non-separable/ state-separable utilities versus time-separable/ state-non-separable utilities is how well they perform empirically. We need a systematic approach to test and discriminate between such competing models. Heaton (1990) reports an interesting study in this direction.

7 References


10. A. Hindy and C. Huang [1990], "Optimal Consumption with Intertemporal Substitution II: The Case of Uncertainty", MIT.


NM preferences over consumption lotteries reflect ordinal preferences over sure consumption and not necessarily risk aversion.

Figure 1  Kihlstrom and Mirman (1974)
Preference Transfer Mapping

Figure 2  Rossman and Selden (1978)
Preference Transfer Mapping
for consumption lotteries

Figure 3    Rossman and Selden (1978)
Coherence Axiom

Figure 4  Rossman and Selden (1978)
Preferences for Intertemporal Consumption

Induced Preferences

Kreps and Porteus (1978)

Primitive Preferences

Coherence Axiom satisfied

Coherence Axiom Violated

Selden (1978)
Epstein and Zin (1989)

Time separable preferences on sure consumption.
Merton (1971)

Non time separable preferences on sure consumption
Hindy and Huang (1990)

Classification of models of intertemporal Consumption

Figure 5