High Hopes and Disappointment

Philip H. Dybvig and L. C. G. Rogers

Preliminary draft: not for redistribution
November 4, 2009

ABSTRACT

We develop a new model of preferences that high expectations about the future make us happy now but make us unhappy later if they are not realized. These preferences generalize the recursive utility of Kreps-Porteus and Epstein-Zinn. Like recursive utility, our agents care about the time resolution of uncertainty, but unlike recursive utility, our agents are not purely forward looking. For example, two agents with the same initial preferences and the same wealth now will feel and behave differently if they are disappointed because they expected to do much better or if they feel lucky because they expected to do much worse. This has some of the flavor of habit formation, but it is different because the effect is there even if there is no intermediate consumption or both agents consumed the same in the past. Agents with these preferences follow path-dependent strategies even when returns are IID, and they tend to behave somewhat cautiously to guarantee the expectations they have formed, making their apparent risk aversion different at different points in time.

1 Introduction.

Recursive Utility models generalize von Neumann-Morgenstern preferences by giving agents preference over the timing of resolution of uncertainty. These preferences have been given an axiomatic foundation by Kreps and Porteus [1978, 1979] and theoretical and empirical development by L. Seldon [1978], Epstein and Zin [1989, 1991], and others. We provide a generalization of these models which does not have the recursive structure. In recursive models, preferences looking forward do not depend on past consumption or beliefs. In our model, preferences looking forward also do not depend on past consumption, but they can depend on past beliefs, or perhaps more generally and more accurately on what was expected to happen on other branches of the tree. Preferences in our model are not defined recursively, which avoids the questions of existence that can be difficult in settings with continuous time and/or an infinite horizon.[1] We use our new model of preferences to derive optimal consumption and investment strategies, which can still be solved using dynamic programming but with an extra state variable.

In our model, there are two contributions to utility, one coming from current consumption and another coming from anticipations of future consumption. In this simplest form, the anticipation depends on the smallest possible future consumption or equivalently there is a constraint that consumption can never be less than the anticipation. Our original motivation for this assumption is the popularity of minimum rates of return in insurance contracts, implying a guarantee on the final value that ratchets up over time. A fixed guarantee (like in portfolio insurance at the investment horizon) can be rationalized using von Neumann-Morgenstern preferences and a subsistence level of consumption, but our model (and some insurance contracts) also allow for a ratcheting up of the minimum guarantee.

Of course, actual 100% sure guarantees are not available in practice, so we have formulated a second model which has "guarantees" that can be violated but at a utility cost. There are different ways of softening a guarantee; our choice is interesting and tractible and allows the agent to choose the referent utility level below which there is a penalty. We can think of this as operationalizing the common expression that "I don't dare to hope" for a good thing. High expectations now make us happy in anticipation but can carry a large psychological cost if they are not met. This penalty from not meeting expectations formed in the past is the new feature that is not present in recursive utility. Although the popular sense of the word "expectations" is a good description of what we are talking about in this paper, it is perhaps not a good term to use in an academic paper because of possible confusion with the mathematical usage of the word "expectations." So, we will call them anticipations instead.

The failure to meet anticipations set in the past give the paper some similarity to preferences with "habit formation," such as Constantinides [1990] or especially Dybvig [1995],[2] because habit formation is often expressed in terms of a reference consumption level such as a previous maximum or weighted-average consumption. However, the reference utilities in the anticipations in this paper are forward-looking and are not tied to past consumption. For example, the preferences in this paper make sense even if consumption takes place only at the end. The referent level of consumption or utility depends on the anticipations we had in the past, not the consumption we had in the past. One manifestation of the difference between the two models is that increasing consumption can only make

[1] See, for example, Duffie and Epstein [1992].

[2] The solution in Dybvig [1995] can be obtained in a limiting case of this paper.

the agents described here better off, while increasing consumption can make an agent with habit formation worse off (because it creates unreasonable demands on consumption in the future).

The prefences we are studying are definitely not recursive, and in particular the choice going forward in a frictionless setting[3] depends not only on wealth and the state variables govening the return process but also on previous anticipations. Alternatively, the anticipations may not be observeable[4] and we can write preferences directly in terms of the process followed by consumption and learning about it (the filtration) if we substitute in the optimal anticipations. In this perspective, preferences can be seen to depend on the distribution of consumption in the whole tree, not just on the continuation of a particular path, which can be viewed of the source of the additional state variables (the past anticipations).

A similar lack of simple recursive structure occurs when optimizing given time-consistent Machina [1982] preferences. These preferences assume that utility depends on the distribution of consumption, but (due to possible failure of the Independence Axiom) may not satisfy the von Neumann-Morgenstern restriction. (Machina[5] preferences are also assumed to satisfy a differentiability property that is important for some analysis but not for our purpose here.) When Machina preferences do not satisfy the Independence Axiom (in the interesting case in which Machina preferences are not also von Neumann Morgenstern preferences), they may seem to be time-inconsistent because applying the same utility function from some node forward can give a different choice than would have been planned at the outset. However, this is probably a mis-application of Machina preferences and Machina preferences are time-consistent if we apply the same preferences computed from the outset in subtrees. (See Machina [1989] who has a nice example with Mom flipping a coin to allocate treats, a less brutal variant of an old example of a man who must decide which of his sons will die.) The distribution of consumption from other subtrees is information that would be missing if we tried to determine choice in a subtree using only the usual forward-looking state variables.

[3]In a model with frictions it may not be possible to change portfolios without additional costs and the state variables may also include such things as individual asset holdings and the distribution of their tax bases.

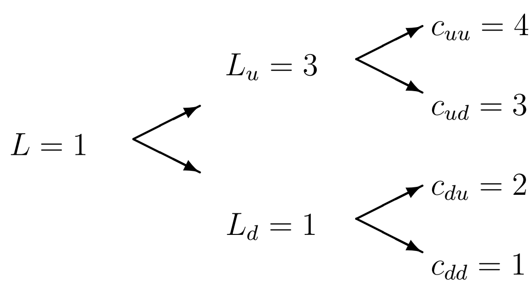
[4]Or, anticipations could be observeable if they can be obtained through survey data or observation of agent characteristics. We do not have a strong view on this either way.

[5]Mark Machina has more than one paper and we should have a subscript ("1" or "1982") to indicate the preferences described in this paper which is best-known to us. Hopefully, this will not lead to significant confusion.

2 Anticipation: Simple Examples

In later sections, we formulate and prove decision problems with anticipation and disappointment in a continuous-time model. In this section, we first want to discuss our model in a simple setting and describe how it is different and more general compared to models with von Neumann-Morgenstern preferences, Machina preferences, or Kreps and Porteus preferences.

The setting for our examples is a binomial model with two intervals (and therefore three points of time) as in Cox, Ross, and Rubinstein (). Therefore, there are four states of nature: $up - up$, $up - down$, $down - up$, and $down - down$, which can be represented in the usual tree diagram. For example, we can write as



a plan (call it Scenario 1) to consume 4 in state $up - up$, 3 in state $up - down$, 2 in state $down - up$, and 1 in state $down - down$. The numbers L, L_u, and L_d, are "anticipations" we will describe shortly and form part of the preferences in our model. In our examples, we will take the state probabilities to be equal (at 1/4 each), which will be convenient for our comparisons.

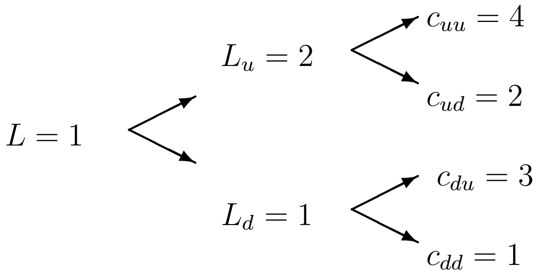
In our model, preferences depend on both *consumption* across states and time (which occurs only at the end in the examples in this section) and *anticipations* of future consumption. Anticipation of high consumption ("high hopes") are a source of happiness, but if unrealized ("disappointment") can be a source of later pain. To model this, we have anticipations at each node before the end, labelled by the moves needed to get there as L, L_u, and L_d. Anticipations cannot decline (we cannot simply adjust our anticipations without psychic cost), and therefore $L_u \geq L$ and $L_d \geq L$. Also (although the model is made more flexible in later sections) we assume now that the agent cannot bear any disappointment and that consumption must always be greater than the anticipation beforehand and therefore $c_{uu} \geq L_u$, $c_{ud} \geq L_u$, $c_{du} \geq L_d$, and $c_{dd} \geq L_d$. Anticipations are choice variables in this model, and we decide how much we dare hope for. There are interesting possibilities that anticipations might be measured separately from consumptions (e.g. through surveys or physical readings of the brain) or manipulated by other agents, and this could possibly lead to models of sentiment or other phenomena. However, in the simple setting in this paper, it matters little whether anticipations are chosen by the agent or chosen by a mechanical rule as a function of the future consumption, and in fact this facilitates the comparison of our model with other models.

Now that we have described consumption and anticipations, we are ready to define preferences, which are represented by a utility function, as follows:

$$(1-p)\left\{ U(L)+\tfrac{1}{2}\left[U(L_u)+U(L_d)\right] \right\}+\tfrac{1}{4}p\left\{ U(c_{uu})+U(c_{ud})+U(c_{du})+U(c_{dd}) \right\}. \quad (1)$$

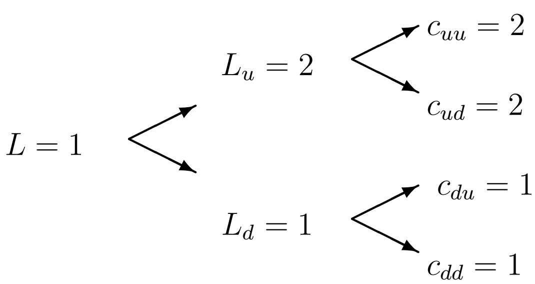
The utility function includes terms for expected utility of terminal consumption as well as expected utility of anticipations. We are not fussing too much in our discrete example about weighting utils at different dates, but we do include a parameter $p \in [0,1]$ that is a measure of how the agent weights utility of consumption relative to utility of anticipation. In practice, we would probably want to assume U is increasing and strictly concave with suitable smoothness, but for the moment we need only assume U is increasing. This will imply that at least some of the constraints we have discussed are binding so that $L_u = \min(c_{uu}, c_{ud})$, $L_d = \min(c_{du}, c_{dd})$, and $L = \min(L_u, L_d) = \min(c_{uu}, c_{ud}, c_{du}, c_{dd})$. If we were thinking of anticipations as being manipulated and doing welfare comparisons based on realizations, we might want to assume something else, but for our purposes in this paper the agent has rational expectations of the law governing information arrival and final consumption, so that we can think of the anticipations as being like nuisance parameters in statistics that can be concentrated and integrated out, or more specifically are functions of the consumptions given above. These rules justify the specifications of L, L_u, and L_d presented in Scenario 1 above.

The purpose of this section is to explore how the preferences represented by the utility function (1) relate to other classes of preferences. Consider first von Neumann-Morgenstern preferences and Machina preferences. Observe that von Neumann-Morgenstern preferences are a limiting case of (1) in which $p = 1$ so there is no weight on anticipations. In Machina preferences, utility is a more general function of the distribution of consumption, subject only to smoothness (a Frechet differentiability condition). In both cases, preferences do not depend on the actual identity of the states in which consumption occurs or on the arrival of infomation about states. Therefore, consider Scenario 2, defined by swapping consumption in the $up-down$ and $down-up$ states of Scenario 1:

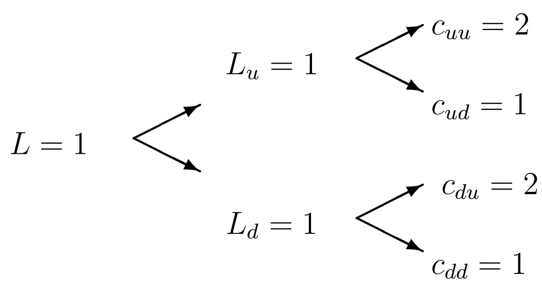


For von Neumann-Morgenstern or Machina preferences, the agent would be indifferent between Scenario 1 and Scenario 2, but our agents prefer Scenario 1 because the anticipation L_u is higher. By monotonicity of U and with the modest additional assumption that U is continuous, we could get the ordering to be opposite for every von Neumann-Morgenstern and Machina agent if we added a sufficiently small amount of consumption in every state to Scenario 2: adding any amount to Scenario 2 makes that strictly more desireable for the von Neumann-Morgenstern and Machina agents, and if the amount is small enough Scenario 2 will still be preferred for our agent.

Perhaps in a general way, the comparison between Scenario 1 and Scenario 2 shows that our preferences are like resolution of uncertainty, like Kreps and Porteus proferences, which is what distinguishes our analysis from von Neumann-Morgenstern or Machina preferences. The following example shows this even more pointedly. Compare Scenario 3,



with Scenario 4,



As before (and throughout the section), we maintain the assumption that the four terminal nodes are equally likely. As a result, Scenario 3 and Scenario 4 each has a 50-50 probability of paying 1 or 2, and von Neumann-Morgenstern preferences or Machina preferences would always be indifferent between the two. However, what is different between the two is the resolution of uncertainty. In Scenario 3, the agent learns the outcome at the end of the first interval of time, and in Scenario 4, the agent learns the outcome only at the very end. Our agents prefers to learn earlier rather than later, and we can see Scenario 3 is preferred to Scenario 4 because because L_u is larger in Scenario 3. As in the comparison between Scenarios 1 and 2, if U is continuous then adding a sufficiently small amount to consumption in all states in Scenario 4 will make the reversal strict.

So far, our examples, have shown how our preferences need not be von Neumann-Morgenstern or Machina preferences. But, are they necessarily Kreps-Porteus preferences? The preferences are actually closely related to Kreps-Porteus preferences, with an additional feature that our preferences about what happens in a subtree may depend on what happens in other contingencies. This is not habit formation (in this section there is no prior consumption to affect subsequent preferences). Rather, if two agents are faced with the same future prospects, an agent who feels they are lucky to be in the current circumstances may have different preferences going forward than an agent who expected better and feels this was a bad draw. In other words, risk preferences may be much different for an agent who expects to be rich but achieves only modest success than for an agent who starts out with very low expectations and achieves the same modest success.

Our example for how our preferences differ from Kreps-Porteus preferences involves four scenarios, and for these scenarios we will assume that $U(1) = 0$, $U(2) = 4$, $U(3) = 7$, and $U(4) = 9$. Consider how these preferences play out for Scenario 5:

$$
L = 1 \quad \begin{matrix} \nearrow \\ \searrow \end{matrix}, \quad \begin{matrix} L_u = 1 \quad \begin{matrix} \nearrow c_{uu} = 4 \\ \searrow c_{ud} = 1 \end{matrix} \\ \\ L_d = 2 \quad \begin{matrix} \nearrow c_{du} = 2 \\ \searrow c_{dd} = 2 \end{matrix} \end{matrix}
$$

Scenario 6:

$$
L = 2 \quad \begin{matrix} \nearrow \\ \searrow \end{matrix}, \quad \begin{matrix} L_u = 2 \quad \begin{matrix} \nearrow c_{uu} = 2 \\ \searrow c_{ud} = 2 \end{matrix} \\ \\ L_d = 2 \quad \begin{matrix} \nearrow c_{du} = 2 \\ \searrow c_{dd} = 2 \end{matrix} \end{matrix}
$$

Scenario 7:

$$
L = 1 \quad \begin{matrix} \nearrow \\ \searrow \end{matrix}, \quad \begin{matrix} L_u = 1 \quad \begin{matrix} \nearrow c_{uu} = 4 \\ \searrow c_{ud} = 1 \end{matrix} \\ \\ L_d = 1 \quad \begin{matrix} \nearrow c_{du} = 1 \\ \searrow c_{dd} = 1 \end{matrix} \end{matrix}
$$

and Scenario 8:

$$
L = 1 \quad \begin{matrix} \nearrow \\ \searrow \end{matrix}. \quad \begin{matrix} L_u = 2 \quad \begin{matrix} \nearrow c_{uu} = 2 \\ \searrow c_{ud} = 2 \end{matrix} \\ \\ L_d = 1 \quad \begin{matrix} \nearrow c_{du} = 1 \\ \searrow c_{dd} = 1 \end{matrix} \end{matrix}
$$

For Kreps-Porteus agents with increasing preferences, the choice between Scenario 5 and Scenario 6 must be the same as the choice between Scenario 7 and Scenario 8, since Scenarios 5 and 6 differ on the same subtree and in the same way as Scenarios 7 and 8. For our agent, the preferences for the two pairs of lotteries depends on p, that is, how important consumption is relative to anticipations. When p is large enough, preferences for consumption dominate and our agent prefers Scenario 7 to Scenario 8 and Scenario 5 to Scenario 6, since for the nodes after *up* that differ, $\frac{1}{2}(U(4)+U(0)) = 4.5 > 4 = \frac{1}{2}(U(2)+U(2))$. For p small enough, preference for anticipations dominate, and our agent prefers Scenario 6 to Scenario 5 and Scenario 8 to Scenario 7, since these scenarios have the larger anticipations (2 instead of 1) more often. In between (and this is the important case), there is a range of p in which our agent prefers Scenario 6 to Scenario 5 but also prefers Scenario 7 to Scenario 8, which cannot happen for Kreps-Porteus agents. This can happen because there is an additional place (the initial node L) where Scenario 6 looks better than Scenario 5 and is not balanced by a corresponding node of Scenario 8 versus Scenario 7. So, if we increase p slightly from the value at which our agent is indifferent between Scenario 7 and Scenario 8, then we have the desired pattern that Scenario 7 is strictly preferred to Scenario 8 but Scenario 6 is strictly preferred to Scenario 5. This cannot happen for Kreps-Porteus agents, who must be have the same preferences over Scenarios 5 and 6 as over Scenarios 7 and 8.

We have shown how our preferences are not in general von Neumann-Morgenstern preferences, Machina preferences, or Kreps-Porteus preferences. We could talk about an axiomatic base for our preferences or a superset of them. For example, we can think of our preferences as being similar to Kreps-Porteus but having an aggregator with more than one variable. But we leave these questions to others and we turn to the question of optimal portfolio choice in a more general version of our model that admits the possibility of a utility penalty but not absolute rigidity about consuming at less than anticipations.

3 Continuous-time Problem and Solution.

Moving to continuous time, there are several modeling choices. We select features that preserve our innovations while implying a stationary model we can solve. It would be possible to have anticipations at each date about every future date, in which case our agent would maximize

$$E\int_{t=0}^{\infty} e^{-\rho t}\{\ (1-p)(\int_{s=t}^{\infty} e^{-\rho(s-t)}U(L_{ts})) + pU(c_t)\ \}\ dt \tag{2}$$

where L_{ts} must be chosen to be nondecreasing in t and $c_s \geq L_{ss}$. However, this model is unwieldy and we prefer the simpler formulation

$$E\int_{0}^{\infty} e^{-\rho t}\{\ (1-p)U(L_t) + pU(c_t)\ \}\ dt \tag{3}$$

maximized subject to the constraint

$$c_t \geq L_t \quad \forall t. \tag{4}$$

In this formulation, the agent has a single anticipation L_t at time t, not a function L_{ts} of the future date s. Note that in continuous time, the contribution of one date to the utility function is null and therefore any distinction between L_t chosen at t and L_t chosen before t is no longer important, even if c_t can have jumps. Therefore we will only require that L_t be nondecreasing, and not fuss about whether it is right-continuous or left-continuous.

In fact, the constraint (4) is arguably too severe, in that there are no certain guarantees in life. We model this by changing the objective to

$$\sup\ E \int_0^\infty e^{-\rho t} \{\ (1-p)U(L_t) + p[\,U(c_t) - K(U(L_t) - U(c_t))^+\,]\ \}\ dt \tag{5}$$

and doing away with the constraint (4). Here, we are assuming that $K \geq 0$ is constant; by taking $K = \infty$ we would find ourselves back in the original situation, so by solving (5) for all positive K including $K = \infty$, we encompass the original problem. We will take $U :\rightarrow (0,\infty) \rightarrow \mathbb{R}$ to be an increasing strictly concave C^2 function satisfying the Inada conditions, and $\rho > 0$ and $0 \leq p \leq 1$ are constant. The value L_0 is given.

For brevity, we shall introduce the notation

$$\begin{aligned} G(c,L) &\equiv (1-p)U(L) + p[\,U(c) - K\{U(L) - U(c)\}^+\,] \qquad (6) \\ &= (1-p)U(L) + p\min\{U(c), (1+K)U(c) - KU(L)\}, \end{aligned}$$

which exhibits G as a concave increasing function of c, with a jump in gradient at $c = L$. Notice that G has no global convexity/concavity property in its second argument. In terms of G, the agent's objective is

$$\sup E \int_0^\infty e^{-\rho t}\, G(c_t, L_t)\ dt. \tag{7}$$

The wealth w_t of the agent at time t evolves from its initial value w_0 according to the conventional dynamics

$$dw_t = rw_t dt + \theta_t(\sigma dZ_t + (\mu - r)dt) - c_t dt, \tag{8}$$

where $r > 0$, $\sigma > 0$ and μ are constants, and Z is a standard Wiener process. To avoid degeneracies, we assume that

$$\mu \neq r. \tag{9}$$

The agent chooses the rate-of-consumption process c, and the portfolio process θ satisfying the non-negativity constraint

$$w_t \geq 0 \quad \forall t \tag{10}$$

on wealth.

To summarize, then, the problem to be solved is the following:

Problem 1: *Given $L \geq 0$ and $w > 0$, obtain*

$$V(w, L) \equiv \sup E\Big[\int_0^\infty e^{-\rho t} G(c_t, L_t)\, dt \;\Big|\; w_0 = w, L_0 = L \Big], \tag{11}$$

where the supremum is taken over all admissible investment/consumption choices (θ, c) which maintain the wealth process w, evolving as

$$dw_t = rw_t dt + \theta_t(\sigma dZ_t + (\mu - r)dt) - c_t dt,$$

nonegative at all times, and is subject to the constraint

L is non-decreasing.

It is necessary to assume that

$$pK > 1 - p. \tag{12}$$

If $pK < 1 - p$, then we take L arbitrarily large, and the problem is ill-posed. If $pK = 1-p$, then by again taking L arbitrarily large we reduce to the conventional separable von Neumann-Morgenstern consumption.

4 Derivation of the Solution

In this section, we state and prove the main result of the paper. The solution is expressed in terms of the state-price density process

$$\xi_t = \exp\{ -(r + \tfrac{1}{2}\kappa^2)t - \kappa Z_t \}, \tag{13}$$

where $\kappa = (\mu - r)/\sigma$ is the market price of risk, and Z is a standard Brownian motion. We also need the notation $K_1 \equiv K + 1$, $k_1 = K_1^{-1}$ in order to state our main result. We impose the standing assumption that the utility U should satisfy[6]:

[6] This condition would have to be satisfied for the basic Merton problem to be well posed.

Assumption U: for all $\lambda > 0$

$$E \int_0^\infty \xi_t I(\lambda e^{\rho t} \xi_t)\, dt < \infty.$$

Theorem 1 *The solution to Problem 1 is characterized by a unique parameter $\lambda > 0$, in terms of which the optimal consumption process c^* and anticipation process L^* are given by*

$$\Lambda_t = p^{-1} \lambda e^{\rho t} \xi_t, \tag{14}$$

$$\underline{\Lambda}_t = \inf_{0 \le s \le t} \Lambda_s, \tag{15}$$

$$\eta_t \equiv \left(\frac{\underline{\Lambda}_t}{z_*} \right) \wedge U'(L_0), \tag{16}$$

$$L_t^* = I(\eta_t), \tag{17}$$

$$c_t^* = \begin{cases} I(\Lambda_t), & \Lambda_t < \eta_t, \\ L_t^*, & \eta_t \le \Lambda_t \le K_1 \eta_t \\ I(k_1 \Lambda_t), & K_1 \eta_t \le \Lambda_t. \end{cases} \tag{18}$$

The constant $z_* \in (0, K_1)$ is obtained by solving an optimal stopping problem for Λ as explained in Propositions 1 and 2 in the Appendix. This optimal stopping problem does not depend on U. The parameter λ is related to the initial wealth w_0 by the budget constraint

$$w_0 = E \int_0^\infty \xi_t c_t^*\, dt, \tag{19}$$

and the value of the problem is

$$V(w_0, L_0) = E \int_0^\infty e^{-\rho t} G(c_t^*, L_t^*)\, dt \tag{20}$$

$$= E \int_0^\infty e^{-\rho t} \tilde{G}(p\Lambda_t, L_t^*)\, dt + \lambda w_0, \tag{21}$$

where $\tilde{G}(y, L) \equiv \sup_x \{G(x, L) - xy\}$ is the convex dual of G.

If

$$p \le p^* \equiv \frac{r}{r + \frac{1}{2}\kappa^2 (1 - K_1^{-b+1})(a+1)} \tag{22}$$

then the agent will never consume in excess of the anticipation level L.

Remarks. (i) The optimal consumption process c^* can be expressed more compactly as

$$U'(c_t^*) = (\eta_t \wedge \Lambda_t) \vee (k_1 \Lambda_t). \tag{23}$$

(ii) Notice that if $z_* > 1$ we have from (16) that $\eta_t < \underline{\Lambda}_t \leq \Lambda_t$, and so the first case in the characterization (18) of c^* never happens. The interpretation here is that the value of p is so high that the agent prefers anticipation so strongly that he pushes the level L up so high that it becomes an aspiration, which may be achieved, but will never be exceeded. Indeed, if his wealth climbs, then he keeps raising the anticipation accordingly, so that it remains always above his actual consumption. This is an interesting and unexpected phenomenon, bearing in mind that we began by thinking of the anticipation as a *lower* bound for consumption.

(iii) Suppose L_0 is fixed, and consider what happens if we decrease λ. As λ falls past the value $\lambda_0 \equiv p z_* U'(L_0)$ we see a qualitative change in (16) ; for $\lambda < \lambda_0$ there is at time 0 an immediate upward jump in the level of L. From the fact that c^* is monotone decreasing in λ, and from (19), we see that *if w_0 exceeds some critical value, there is at time 0 an immediate raising of the anticipation level L.*

PROOF. The proof proceeds through a sequence of stages, which are

1. Set up the Lagrangian form of the problem;
2. Optimize over c, assuming L given;
3. Optimize over increasing adapted L;
4. Confirm that the candidate solution is the optimal solution.

1. SET UP THE LAGRANGIAN FORM OF THE PROBLEM. As the market is complete, the agent may fund any consumption stream c which satisfies the budget constraint

$$E \int_0^\infty \xi_t c_t \, dt = w_0. \tag{24}$$

We may therefore cast the agent's problem in Lagrangian form as

$$\sup_{c,L} \Psi(c, L; \lambda) \equiv \sup_{c,L} \; E \int_0^\infty e^{-\rho t} \{ \; G(c_t, L_t) - \lambda e^{\rho t} \xi_t c_t \; \} \, dt + \lambda w_0. \tag{25}$$

We introduce the notation

$$\Lambda_t \equiv p^{-1} \lambda e^{\rho t} \xi_t \tag{26}$$

in terms of which the Lagrangian optimization problem is expressed as

$$\sup_{c,L} \; E \int_0^\infty e^{-\rho t} \{ \; G(c_t, L_t) - p \Lambda_t c_t \; \} \, dt + \lambda w_0. \tag{27}$$

Notice particularly that $\xi_0 = 1$, so $\Lambda_0 = \lambda/p$.

2. OPTIMIZING OVER c. The terms involving c can easily be maximised to give the first-order condition

$$U'(c_t)(1 + KI_{\{L_t > c_t\}}) = \Lambda_t. \tag{28}$$

Writing I for the inverse to U', we can characterise the optimal consumption quite explicitly in terms of the multiplier process Λ, and the constants $K_1 \equiv K + 1$, $k_1 \equiv K_1^{-1}$:

$$c_t^* = \begin{cases} I(\Lambda_t), & \Lambda_t < U'(L_t); \\ L_t, & U'(L_t) \le \Lambda_t \le K_1 U'(L_t); \\ I(k_1 \Lambda_t), & K_1 U'(L_t) < \Lambda_t. \end{cases} \tag{29}$$

This establishes the form (18) of c^*, and we only need to find out what L^* should be. The optimised value for the terms involving c is

$$\begin{aligned} \tilde{G}(p\Lambda_t, L_t) &= G(c_t^*, L_t) - p\Lambda_t c_t^* \\ &= (1-p)U(L_t) + p \begin{cases} \tilde{U}(\Lambda_t), & \Lambda_t < U'(L_t); \\ U(L_t) - L_t \Lambda_t, & U'(L_t) \le \Lambda_t \le K_1 U'(L_t); \\ K_1 \tilde{U}(k_1 \Lambda_t) - KU(L_t), & \Lambda_t > K_1 U'(L_t). \end{cases} \end{aligned}$$

3. OPTIMIZING OVER L. The above expression for $\tilde{G}(p\Lambda_t, L_t)$ depends on L_t, a variable that we have yet to optimise over. We shall turn the problem of optimizing over the increasing process L into an optimal stopping problem for the state-price density process, which turns out to be more amenable.

To do this, we need to express $\tilde{G}(p\Lambda, L)$ as an integral; differentiating $\tilde{G}(p\Lambda, L)$ with respect to L leads to the expression

$$\frac{\partial \tilde{G}}{\partial L}(p\Lambda, L) - (1-p)U'(L) = p \begin{cases} 0, & \Lambda < U'(L); \\ U'(L) - \Lambda, & \Lambda \le K_1 U'(L); \\ -KU'(L), & \Lambda > K_1 U'(L). \end{cases} \tag{30}$$

This is more compactly expressed as

$$\begin{aligned} \frac{\partial \tilde{G}}{\partial L}(p\Lambda, L) &= (1-p)U'(L) - p\big\{ (\Lambda - U'(L))^+ \wedge KU'(L) \big\} \\ &= U'(L)\left\{1 - p - p\left[\left(\frac{\Lambda}{U'(L)} - 1\right)^+ \wedge K \right]\right\} \end{aligned} \tag{31}$$

The objective to be maximized over increasing L can therefore be written as

$$\begin{aligned} Q &\equiv E\int_0^\infty e^{-\rho t}\tilde{G}(p\Lambda_t, L_t)\,dt \\ &= E\int_0^\infty e^{-\rho t}\left\{\tilde{G}(p\Lambda_t, L_0) + \int_{L_0}^{L_t}\frac{\partial\tilde{G}}{\partial L}(p\Lambda_t, x)dx\right\}dt \\ &= E\int_0^\infty e^{-\rho t}\tilde{G}(p\Lambda_t, L_0)\,dt + \\ &\qquad +E\int_0^\infty e^{-\rho t}\int_{L_0}^\infty I_{\{x\le L_t\}}\frac{\partial\tilde{G}}{\partial L}(p\Lambda_t, x)\,dx\,dt \\ &\equiv Q_0 + Q_1, \end{aligned}$$

say. Notice that the first term $Q_0 = E\int_0^\infty e^{-\rho t}\tilde{G}(p\Lambda_t, L_0)dt$ is determined solely by Λ_0, and L_0, and does not depend in any way on the choice of L_t for $t > 0$.

We now define

$$f(y) \equiv E\Big[\int_0^\infty e^{-\rho t}\{\, 1 - p - p[\, (\Lambda_t - 1)^+ \wedge K\,]\,\}\,dt\,\Big|\,\Lambda_0 = y\,\Big], \tag{32}$$

a function whose form is made explicit in Proposition 1. We are able to express the objective Q_1 in terms of f as follows:

$$\begin{aligned} Q_1 &= E\int_0^\infty e^{-\rho t}\int_{L_0}^\infty I_{\{x\le L_t\}}\frac{\partial\tilde{G}}{\partial L}(p\Lambda_t, x)\,dx\,dt & (33)\\ &= \int_{L_0}^\infty E\Big[\int_{\tau_x}^\infty e^{-\rho t}\frac{\partial\tilde{G}}{\partial L}(p\Lambda_t, x)\,dt\Big]\,dx & (34)\\ &= \int_{L_0}^\infty U'(x)E\big[e^{-\rho\tau_x}f(\Lambda_{\tau_x}/U'(x))\big]\,dx. & (35) \end{aligned}$$

We now consider the maximization of Q_1 over non-decreasing L as a family of optimal stopping problems, $\sup_\tau E[e^{-\rho\tau}f(\Lambda_\tau/U'(x))]$, one for each x[7]. However, because of the exponential form of the diffusion Λ, the mutiplicative factor $U'(x)$ can be absorbed into the initial condition for Λ, and there is in reality *just one* optimal stopping problem, which is to find

$$\bar{f}(y) \equiv \sup_\tau E[e^{-\rho\tau}f(\Lambda_\tau)\,|\,\Lambda_0 = y\,]. \tag{36}$$

We shall see (Proposition 2) that the optimal stopping time for this problem is of the form

$$\tau^* \equiv \inf\{t : \Lambda_t \le z_*\} \tag{37}$$

[7]There is no guarantee that the optimal stopping times τ_x will increase with x - which would be a problem since we want them to be inverse to the increasing process L. However, it will transpire that they do indeed have this property.

for some positive constant $z_* \in (0, K_1)$ which is the unique root of (A10). Thus the optimal stopping time τ_x will be simply

$$\tau_x \equiv \inf\{t : \Lambda_t \le z_* U'(x)\}, \tag{38}$$

from which it is immediate that *the stopping times* τ_x *increase with* x. Hence from the (almost-sure) equality of events[8] for each $x > L_0$ the optimal guarantee level L^* must satisfy

$$\{L^*_t > x\} = \{\tau_x < t\} = \{\underline{\Lambda}_t < z_* U'(x)\} = \{I(z_*^{-1}\underline{\Lambda}_t) > x\}.$$

we learn that the constructed L is *increasing*, and may even be expressed explicitly as

$$L^*_t = \max\{\, L_0, I(z_*^{-1}\underline{\Lambda}_t)\,\} = I\left(\frac{\underline{\Lambda}_t}{z_*} \wedge U'(L_0)\right). \tag{39}$$

This establishes the state form (17) of L^*.

We therefore have that

$$\begin{aligned} Q_1 &= \int_{L_0}^{\infty} U'(x)\bar{f}(\Lambda_0/U'(x))\, dx & (40)\\ &= -\Lambda_0^2 \int_{\Lambda_0/U'(L_0)}^{\infty} \frac{\bar{f}(y)}{y^3 U''(I(\Lambda_0/y))}\, dy & (41)\\ &= \Lambda_0^2 \int_{\Lambda_0/U'(L_0)}^{\infty} y^{-3}\tilde{U}''(\Lambda_0/y)\bar{f}(y)\, dy, & (42)\end{aligned}$$

after some routine calculations.

4. Confirming the optimality of the constructed solution. We have now constructed what we believe is the optimal solution to the problem, in terms of the state-price density process ξ. In fact, the optimal consumption process given by (29) is only defined up to the scalar multiple λ. Notice that as λ increases, the consumption process $c^* \equiv c^*(\lambda)$ defined at (29) decreases, and that always

$$I(\Lambda_t) \le c^*_t \le I(k_1\Lambda_t). \tag{43}$$

Given Assumption U, the map

$$\lambda \mapsto E\int_0^\infty \xi_t c^*_t(\lambda)\, dt$$

is continuous and monotone decreasing (by monotone convergence). Since we have also assumed the Inada conditions, the map is onto $(0,\infty)$, so for any $w_0 > 0$ there is a unique value $\lambda_*(w_0)$ such that

$$E\int_0^\infty \xi_t c^*_t(\lambda_*(w_0))\, dt = w_0. \tag{44}$$

[8]We use the notation $\underline{\Lambda}_t \equiv \inf_{0\le s\le t}\Lambda_s$.

Thus for given $w_0 > 0$, the consumption process $c^*(\lambda_*(w_0))$ satisfies the budget constraint (24) and so is feasible. Moreover, by construction it maximises the Lagrangian $\Psi(c, L; \lambda_*(w_0))$ and therefore is optimal.

□

5 Solution for CRRA utility.

Section 4 shows how Problem 1 can be solved for a quite general C^2 strictly concave utility U, leading to an optimal solution which is characterized in terms of various integrals which can only be evaluated once U is specified explicitly[9]. This section is devoted to the special case of CRRA utility:

$$U(x) = \frac{x^{1-R}}{1-R}, \tag{45}$$

where the coefficient of relative risk aversion, R, is a positive constant different from 1. With this simplifying assumption, we are able to characterize the value function V through its convex dual, and obtain a remarkably complete solution which is readily computed. To state the solution, we need the notations:

$$\tilde{U}(z) \equiv -\frac{z^{1-R'}}{1-R'}, \tag{46}$$

for the convex dual of U, where

$$R' \equiv R^{-1}; \tag{47}$$

$-a < 0$ and $b > 1$ for the two roots of the quadratic

$$Q(t) \equiv \tfrac{1}{2}\kappa^2 t(t-1) + (\rho - r)t - \rho; \tag{48}$$

and

$$K_1 \equiv K + 1, \quad \gamma_M \equiv R^{-1}\{\rho + (R-1)(r + \kappa^2/2R)\}. \tag{49}$$

The constant γ_M is the proportional rate of consumption in the conventional Merton problem.

Theorem 2 *The value function V satisfies the scaling relation*

$$V(w, L) = L^{1-R} V(wL^{-1}, 1) \equiv L^{1-R} v(wL^{-1}). \tag{50}$$

[9] This is no different from the (much simpler) situation where $K = 0$ and $p = 1$, the familiar Merton problem.

The function v is concave and increasing, and is characterized in terms of its convex dual function

$$J(z) \equiv \sup\{v(x) - zx\}, \tag{51}$$

as follows.

(i) If $p > p^*$, then for some $z_0 \in (0, p)$ and constants C, A_0, B_0, A_1, B_1, and A_2,

$$J(z) = \begin{cases} C\tilde{U}(z) & (z \leq z_0) \\ \rho^{-1}(1-p)U(1) + p^{1/R}\gamma_M^{-1}\tilde{U}(z) + A_0 z^{-a} + B_0 z^b & (z_0 \leq z \leq p) \\ \rho^{-1}U(1) - z/r + A_1 z^{-a} + B_1 z^b & (p \leq z \leq K_1 p) \\ \rho^{-1}(1-K_1 p)U(1) + (K_1 p)^{1/R}\gamma_M^{-1}\tilde{U}(z) + A_2 z^{-a} & (K_1 p \leq z) \end{cases}$$

where we require J to be C^1 everywhere, and to be C^2 at z_0.

(ii) If $p < p^*$, then for some $z_0 \in (p, K_1 p)$ and constants C, A_1, B_1, and A_2,

$$J(z) = \begin{cases} C\tilde{U}(z) & (z \leq z_0) \\ \rho^{-1}U(1) - z/r + A_1 z^{-a} + B_1 z^b & (z_0 \leq z \leq K_1 p) \\ \rho^{-1}(1-K_1 p)U(1) + (K_1 p)^{1/R}\gamma_M^{-1}\tilde{U}(z) + A_2 z^{-a} & (K_1 p \leq z) \end{cases}$$

where we require J to be C^1 everywhere, and to be C^2 at z_0.

The value of p^* is given by the expression at (22).

Remarks. (i) In the first case of the Theorem, we have a large value of p, therefore relatively little weight on the utility of a high guarantee level. Optimal consumption may from time to time fall below L; it may stick at L for quite long periods of time, and it may also be in excess of L. If wealth is high enough relative to the guarantee level L, the agent will raise L, but there will be an intermediate region in which the agent consumes above L yet wealth is not high enough to justify raising the guarantee level.

(ii) In the second case of the Theorem, the weight on the utility of the guarantee level is higher, and it is in the agent's best interests to get L as high as possible. As for the first case, it may happen from time to time that consumption falls below the guarantee level, but in this instance we will never consume at rate greater than L; as wealth rises, we may gradually raise L, but the consumption will not exceed L. This qualitative behaviour is strange when we view the guarantee level as some sort of lower bound for future consumption, yet we find that consumption is always below the guarantee level!

(iii) In the first instance, there are seven unknowns to find, z_0, C, A_0, B_0, A_1, B_1, and A_2, and seven conditions to fix them (C^2 at z_0 gives three, C^2 at p gives two, C^1 at $K_1 p$ gives two). In the second instance, there are five unknowns, z_0, C, A_1, B_1, and A_2 to find, and five conditions.

(iv) In the proof to follow, we shall derive the Hamilton-Jacobi-Bellman (HJB) equations for the value of the problem, and solve them quite explicitly. We are well aware that constructing a solution to the HJB equations does not prove optimality, as the argument involved only proves that the value process is a local martingale (not a martingale) under the conjectured optimal control. A verification result is required. We shall omit this, for various reasons. Firstly, the results of Section 4 can in any case be used to demonstrate optimality. Secondly, the HJB approach used here is easier to understand and present than the sample-path approach of Section 4. Thirdly, a verification result would be lengthy and technical, and would not increase understanding.

PROOF. The scaling assertion (50) follows from a familiar argument. Invoking the Martingale Principle of Optimal Control, we require that the value process

$$Y_t \equiv \int_0^t e^{-\rho t} G(c_t, L_t)\, dt + e^{-\rho t} V(w_t, L_t) \quad \text{is a supermartingale} \tag{52}$$

for any admissible controls, and a martingale for optimal control. Making an Itô expansion of Y, we obtain

$$e^{\rho t} dY_t \doteq \Big\{ G(c,L) - \rho V + (rw + \theta(\mu - r) - c)V_w + \tfrac{1}{2}\sigma^2\theta^2 V_{ww} \Big\}\, dt + V_L\, dL_t, \tag{53}$$

where the notation $\doteq$ signifies that the two sides differ by the differential of a local martingale. The HJB equations for this problem are therefore

$$0 \geq \sup_{c,\theta} \big[G(c,L) - \rho V + (rw + \theta(\mu - r) - c)V_w + \tfrac{1}{2}\sigma^2\theta^2 V_{ww} \big], \tag{54}$$

$$0 \geq V_L \tag{55}$$

with equality in at least one of (54), (55) everywhere. Thus by exploiting the scaling relation (50) the HJB equations (54), (55) can be written as

$$\max \Big\{ \sup_{s,\theta} \mathcal{G}v(x,s,\theta), (1-R)v(x) - xv'(x) \Big\} = 0, \tag{56}$$

where

$$\mathcal{G}v(x,s,\theta) \equiv G(s,1) - \rho v(x) + (rx + \theta(\mu - r) - s)v'(x) + \tfrac{1}{2}\sigma^2\theta^2 v''(x). \tag{57}$$

The optimization of $\mathcal{G}v(x,s,\theta)$ in (67) is easily achieved[10], and tells us that in any region where this inequality holds with equality we shall have

$$0 = \tilde{G}(v',1) - \rho v + rxv' - \tfrac{1}{2}\kappa^2 \frac{(v')^2}{v''}, \tag{58}$$

[10]Notice that the supremum over θ in (54) will be infinite unless $V_{ww} < 0$ everywhere, which implies that V must be strictly concave in w.

where as usual $\tilde{G}$ denotes the convex dual of G,

$$\tilde{G}(z,L) \equiv \sup_x \big[\, G(x,L) - xz \,\big]. \tag{59}$$

We need to make $\tilde{G}$ more explicit, using the definition (6). We have that

$$G_c(c,1) = \begin{cases} pc^{-R} & (c > 1) \\ pK_1 c^{-R} & (c \leq 1). \end{cases}$$

By looking at Figure 1, we see that the inverse to G_c is (recalling that $R' = R^{-1}$)

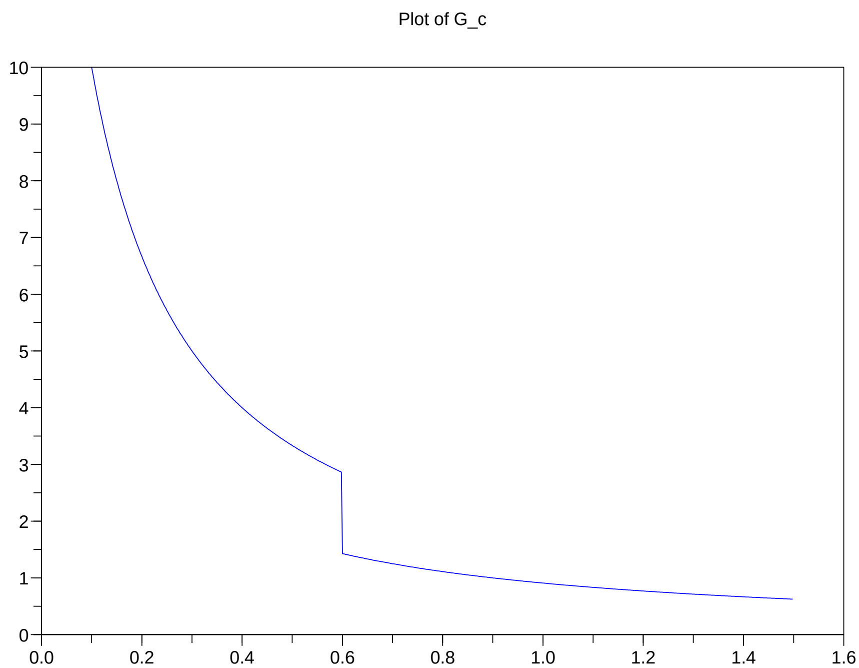


Figure 1: Plot of the derivative G_c of G.

$$g(y,1) = \begin{cases} (p/y)^{R'} & (y \leq p) \\ 1 & (p < y < pK_1) \\ (pK_1/y)^{R'} & (pK_1 \leq y). \end{cases}$$

After some further routine calculations we learn that

$$\tilde{G}(y,1) - (1-p)U(1) = \begin{cases} y^{1-R'} p^{R'} R/(1-R) & (y \leq p) \\ pU(1) - y & (p < y < pK_1) \\ y^{1-R'} (pK_1)^{R'} R/(1-R) & (pK_1 \leq y). \end{cases}$$

More simply,

$$\tilde{g}(z) \equiv \tilde{G}(z,1) = \begin{cases} (1-p)U(1) + p\tilde{U}(z/p) & (z \le p) \\ U(1) - z & (p \le z \le K_1 p) \\ (1-pK_1)U(1) + pK_1\tilde{U}(z/pK_1) & (K_1 p \le z). \end{cases} \tag{60}$$

Thus in the regions where (58) holds with equality, we shall have (see (58))

$$\rho v - r x v' + \kappa^2 \frac{(v')^2}{2v''} = \tilde{g}(v'). \tag{61}$$

The other part of the HJB equations, (55), applies with equality in regions where it is optimal to raise the guarantee level L, which we expect will be regions where the wealth is high relative to the guarantee level. In terms of the rescaled value v, (55) says

$$(1-R)v(x) - xv'(x) \le 0. \tag{62}$$

In view of the remarks made after Theorem 1, we know that for a fixed L_0 there is a threshold value of w_0 such that if w_0 exceeds that value then the optimal behaviour is to raise the anticipation level immediately. In currrent context, what this means is that *the region where* (62) *holds with equality is a half-line* (x_0, ∞)*.*

The next step is to introduce a change of variables, writing

$$z \equiv v'(x), \quad J(z) = v(x) - xz. \tag{63}$$

As is well known, the identities $J' = -x$, and $J'' = -1/v''$ hold, reducing the differential operator to *linear* form; the equation (61) becomes

$$\mathcal{L}J(z) + \tilde{g}(z) = 0 \tag{64}$$

where

$$\mathcal{L} \equiv \tfrac{1}{2}\kappa^2 z^2 \frac{d^2}{dz^2} + (\rho - r)z \frac{d}{dz} - \rho. \tag{65}$$

Since U is known explicitly, this allows us to solve the ODE for J piecewise; the only issue to be discussed is the boundary conditions at the edges of the different regions, which we address shortly. But firstly we observe that in view of Assumption A, there is some $A > 0$ such that $v(x) = AU(x)$ for all $x \ge x_0$, and hence we conclude that the dual function J must be of the form

$$J(z) = C\tilde{U}(z) \tag{66}$$

for all $z \le z_0 \equiv v'(x_0)$, for some constant $C > 0$. To understand the boundary conditions at x_0, firstly note that in view of the scaling, the HJB equations (54), (55) can be written as

$$\max\left\{(1-R)v(x) - xv'(x), \sup_{c,\theta} \mathcal{G}v(x,c,\theta)\right\} = 0, \tag{67}$$

where

$$\mathcal{G}v(x,c,\theta) \equiv G(c,1) - \rho v(x) + (rx + \theta(\mu - r) - c)v'(x) + \tfrac{1}{2}\sigma^2\theta^2 v''(x). \qquad (68)$$

Now observe that v must be C^1, for if the value function had a jump in gradient[11] at some point x_*, then the Itô expansion of $V(w_t, L_t)$ would have a term involving the local time at x_*, and therefore could not be a martingale.

Next, we consider the possible change of v'' at x_0. From (68), we see that $v''(x_0+) \leq v''(x_0-)$, else the supremum (67) on the right of x_0 would be positive. However, the fact that $q(x) \equiv (1-R)v(x) - xv'(x)$ is everywhere non-positive, and equals 0 to the right of x_0 implies that the gradient of q at x_0- must be non-negative, and hence that $v''(x_0-) \leq v''(x_0+)$. Putting these together, we learn that v must be C^2 at x_0, and hence that J must be C^2 at z_0.

This has implications for the joining of (66) to the solution of (64). Since $\tilde{g}$ is defined piecewise, the solution to (64) is also defined piecewise, and solving this equation gives the generic solution

$$J(z) = \begin{cases} \rho^{-1}(1-p)U(1) + p^{1/R}\gamma_M^{-1}\tilde{U}(z) + A_0 z^{-a} + B_0 z^b & (0 < z \leq p) \\ \rho^{-1}U(1) - z/r + A_1 z^{-a} + B_1 z^b & (p \leq z \leq K_1 p) \\ \rho^{-1}(1-K_1 p)U(1) + (K_1 p)^{1/R}\gamma_M^{-1}\tilde{U}(z) + A_2 z^{-a} & (K_1 p \leq z) \end{cases} \qquad (69)$$

for some constants A_0, B_0, A_1, B_1 and A_2. Although the solution to (64) should in principle have a further term in z^b to the right of $K_1 p$, this can be excluded as we know that J must be convex non-increasing, and this property would be destroyed[12] by such a term in the solution.

The constants in the solution are constrained by the C^1 condition, and by the fact that J must be C^2 at z_0. While it is possible to have a C^2 join of (69) to some multiple of $\tilde{U}$ at $z_0 \leq K_1 p$, it is not possible for any $z_0 > K_1 p$. To understand why, consider the difference between the two solutions, which has the form

$$\varphi(z) = -\beta + \alpha\tilde{U}(z) + A_2 z^{-a} \; ;$$

this has to have a C^2 join to zero at z_0, where $\beta = \rho^{-1}(pK_1 - 1)$, and $A_2 > 0$. From this, we derive the conditions

$$\begin{aligned} A_2 z_0^{-a} + \alpha\tilde{U}(z) &= \beta, \\ -aA_2 z^{-a} + (1-R')\alpha\tilde{U}(z) &= 0, \\ a(a+1)A_2 z^{-a} - R'(1-R')\alpha\tilde{U}(z) &= 0. \end{aligned}$$

[11] As we have seen, v must be concave, so the gradient must be monotone decreasing.

[12] Recall that $b > 1$, so that a term in z^b to the right of $K_1 p$ would be the dominant term, and would ultimately make the function either non-decreasing, or not convex.

From the last two of these, we see that $R' = a+1$, so that the function φ is of the form $-\beta + \alpha\tilde{U}(z)$. It is obvious that this cannot have a C^2 join to zero anywhere.

Thus there are just the two possibilities, described in alternatives (i) and (ii) in the statement of the Theorem. By considering the case of C^2 contact at $z_0 = p$, we discover after some calculations that the changeover from $z_0 < p$ to $z_0 > p$ occurs when $p = p^*$ as defined at (22).

□

5.1 Numerical examples.

We shall illustrate the preceding results with some numerical examples. The plots produced show for various parameter settings:

- the dual value function, $\log((1-R)V)$, the optimal proportion of wealth to be held in the stock, and the optimal consumption rate as a function of wealth;
- Simulated sample paths of consumption and wealth.

The requisite calculations for the first set of plots have been dealt with above, but there are still some details to be elucidated concerning the second set of plots, notably the relationship between the initial wealth w_0, and the Lagrange multiplier λ and initial value L_0 of the anticipation. We explain in more detail how this is done in A.3; the full calculations are available on request.

We shall suppose throughout that $\sigma = 0.2$, $\mu = 0.12$, $r = 0.05$, $\rho = 0.1$, $R = 2$, and that $L_0 = 1$. We fix the value of p to be 0.5, and consider what happens for different values of $K = 1.03, 2.03, 3.03, 43$. In the first two examples, the anticipation level is an aspiration, and consumption never exceeds L, whereas in the last two the anticipation level is in general a lower bound for consumption, though it may exceed consumption if things go badly. For the first case, $K = 1.03$, the penalty for falling short of the anticipation level is not very large, so we see in the simulated paths Figure 3 that even though the anticipation level is raised quite a bit in the first half of the plot, the consumption frequently falls below it. In the second half of the plot, when the wealth falls, there is a corresponding fall-off in consumption. The proportion of wealth invested in the risky asset (see Figure 2) varies little with wealth. When K is raised to 2.03, the variation of proportion of wealth in the risky asset is much more pronounced. As we raise K to 3.03, this variation gets even more marked, and as expected, we start to see (in Figure 7) some consumption in excess of the anticipation. Pushing K up to 43 (see Figures 8 and 9) induces a very wide range of proportion of wealth invested in the risky asset, and we begin to see the anticipation level behave as a lower bound for consumption; indeed, in Figure 9 we see that even when the wealth drops away in the second half of the time period, the consumption does not fall below the anticipation level, consistent with the 'guarantee' interpretation of L

A Calculations

A.1 Finding f.

Here we carry out some calculations needed earlier in the paper. At (32), we defined

$$f(y) \equiv E\Big[\int_0^\infty e^{-\rho t} \{ \, 1 - p - p[\, (\Lambda_t - 1)^+ \wedge K \,] \, \} \, dt \, \Big| \, \Lambda_0 = y \Big].$$

Proposition 1 *The function f is given explicitly as*

$$f(x) = \begin{cases} \rho^{-1}(1-p) \, + B_1 x^b, & 0 \le x \le 1 \\ \rho^{-1} - \, px/r \, + A_0 x^{-a} + B_0 x^b, & 1 \le x \le K_1 \\ \rho^{-1}(1 - pK_1) + A_1 x^{-a}, & K_1 \le x, \end{cases} \tag{A1}$$

where $-a < 0$ and $b > 1$ are the roots of the quadratic

$$Q(t) \equiv \tfrac{1}{2}\kappa^2 t(t-1) + (\rho - r)t - \rho, \tag{A2}$$

and

$$\rho r(a+b) \begin{pmatrix} A_0 \\ B_0 \\ A_1 \\ B_1 \end{pmatrix} = -\tfrac{1}{2}p\kappa^2 \begin{pmatrix} b(b-1) \\ -K_1^{-b+1} a(a+1) \\ -(K_1^{a+1} - 1) b(b-1) \\ (1 - K_1^{-b+1}) a(a+1) \end{pmatrix}. \tag{A3}$$

Proof. By Itô's formula, we see that f must satisfy

$$1 - p - p[\, (x-1)^+ \wedge K \,] - \rho f(x) + \tfrac{1}{2}\kappa^2 x^2 f''(x) + (\rho - r) x f'(x) = 0, \tag{A4}$$

with the conditions that f is C^1 at 1 and $K_1 \equiv K + 1$. The homogeneous differential equation has solutions x^{-a}, x^b, where $-a < 0 < 1 < b$ are the roots of the quadratic equation

$$\tfrac{1}{2}\kappa^2 t(t-1) + (\rho - r)t - \rho = 0. \tag{A5}$$

Particular solutions of the inhomogeneous differential equation are easy to find by inspection. Since f must remain bounded at 0 and ∞, we deduce that f is of the form

$$\begin{aligned} f(x) &= \frac{1-p}{\rho} + B_1 x^b & & (0 \le x \le 1) \\ &= -\frac{px}{r} + \frac{1}{\rho} + A_0 x^{-a} + B_0 x^b & & (1 \le x \le K_1) \\ &= \frac{1 - pK_1}{\rho} + A_1 x^{-a} & & (K_1 \le x) \end{aligned}$$

The C^1 conditions give four linear equations for the coefficients:

$$\begin{pmatrix} 1 & 1 & 0 & -1 \\ -a & b & 0 & -b \\ K_1^{-a} & K_1^b & -K_1^{-a} & 0 \\ -aK_1^{-a} & bK_1^b & aK_1^{-a} & 0 \end{pmatrix} \begin{pmatrix} A_0 \\ B_0 \\ A_1 \\ B_1 \end{pmatrix} = \frac{p}{\rho r} \begin{pmatrix} \rho - r \\ \rho \\ K_1(\rho - r) \\ K_1 \rho \end{pmatrix}$$

Routine calculations lead us to the solution:

$$\begin{aligned} \rho r(a+b) \begin{pmatrix} A_0 \\ B_0 \\ A_1 \\ B_1 \end{pmatrix} &= -p \begin{pmatrix} br - (b-1)\rho \\ K_1^{-b+1}(ar - (a+1)\rho) \\ (K_1^{a+1} - 1)((b-1)\rho - br) \\ (K_1^{-b+1} - 1)(ar - (a+1)\rho) \end{pmatrix} \\ &= -\tfrac{1}{2} p \kappa^2 \begin{pmatrix} b(b-1) \\ -K_1^{-b+1} a(a+1) \\ -(K_1^{a+1} - 1) b(b-1) \\ (1 - K_1^{-b+1}) a(a+1) \end{pmatrix} \end{aligned}$$

Observe that $A_0, B_1 < 0$ and $A_1, B_0 > 0$. Notice also from (A4) that f is C^2, and from the form of the solution we see also that f' must be unimodal.

A.2 Solving the optimal stopping problem.

We here discuss the solution of the optimal stopping problem (36):

$$\bar{f}(y) \equiv \sup_{\tau} \; E\big[e^{-\rho\tau} f(\Lambda_\tau) \;\big|\; \Lambda_0 = y\,\big], \tag{A6}$$

where f is as defined at (32), made explicit in Proposition 1. We find the following result.

Proposition 2 *The optimal solution to the stopping problem* (A6) *is given by taking*

$$\tau = \tau^* \equiv \inf\{t : \Lambda_t \leq z_*\}, \tag{A7}$$

where z_* is the unique solution to $H(z) = 0$, where H is defined as

$$\rho r H(z) = \begin{cases} ra(1-p) - \frac{1}{2}p\kappa^2(1 - K_1^{-b+1})a(a+1)z^b, & 0 \leq z \leq 1 \\ ra - \rho(a+1)pz + \frac{1}{2}p\kappa^2 K_1^{-b+1} a(a+1) z^b, & 1 \leq z \leq K_1 \\ ra(1 - pK_1), & K_1 \leq z. \end{cases} \tag{A8}$$

The solution z_* is greater than 1 if and only if

$$p < p^* \equiv \frac{r}{r + \frac{1}{2}\kappa^2(1 - K_1^{-b+1})(a+1)}. \tag{A9}$$

Proof. Evidently f is decreasing, with limit $(1-p-pK)/\rho < 0$ at infinity. We shall also need to make use of the fact that

$$f = R_\rho h,$$

where $h(x) = 1 - p - p((x-1)^+ \wedge K)$ is decreasing, and R_ρ is the ρ-resolvent operator[13] of the diffusion Λ.

The solution to the optimal stopping problem satisfies

$$\max\{(\mathcal{G}-\rho)\bar{f}, f - \bar{f}\} = 0$$

In the stopping region, where $\bar{f} = f$, we have

$$(\mathcal{G}-\rho)\bar{f} = (\mathcal{G}-\rho)f = -h,$$

so the stopping region is contained entirely in $F_0 \equiv \{x : h(x) \geq 0\} = [0, p^{-1}] \subseteq [0, K_1]$. Is it possible that there might be an interval $F_c = (\alpha, \beta) \subset F_0$ such that the optimal policy was to continue in F_c and stop at the endpoints? If this happened, then certainly $h > 0$ throughout I_c, and so

$$f(x) = E^x\Big[\int_0^\tau e^{-\rho t} h(\Lambda_t)\,dt + e^{-\rho\tau} f(\Lambda_\tau)\Big] > E^x\Big[e^{-\rho\tau} f(\Lambda_\tau)\Big] = E^x\Big[e^{-\rho\tau}\bar{f}(\Lambda_\tau)\Big]$$

where $\tau = \inf\{t : \Lambda_t \notin (\alpha, \beta)\}$. This would mean it was better to stop immediately in (α, β) rather than wait til the process gets out of that interval, a contradiction.

The optimal rule is therefore of the form $\tau = \inf\{t : \Lambda_t \leq z_*\}$ as claimed, and it remains only to find the critical value of z_*. The solution $\bar{f}$ is of the form

$$\begin{aligned} \bar{f}(y) &= Ay^{-a} && (y \geq z_*) \\ &= f(y) && (y \leq z_*). \end{aligned}$$

Continuity of $\bar{f}$ at z_* gives that $\bar{f}(y) = f(z_*)(y/z_*)^{-a}$ for $y \geq z_*$, and so we see[14] that we must choose z_* to maximize $z^a f(z)$; that is, we shall require to solve the equation

$$H(z) \equiv af(z) + zf'(z) = 0 \tag{A10}$$

in order to locate z_*. The reader can easily check that the form given at (A8) for H is what develops from the identity $H(z) = af(z) + zf'(z)$ when f is as given by Proposition 1.

[13] Formally, $R_\rho = (\rho - \mathcal{G})^{-1}$, where $\mathcal{G} \equiv \frac{1}{2}\kappa^2 x^2 \frac{d^2}{dx^2} + (\rho - r)x\frac{d}{dx}$ is the generator of Λ.

[14] Hold $y \geq p^{-1}$ fixed ...

Notice that H is clearly decreasing in $[0,1]$, with $H(0) > 0$. We also see that H is constant and negative in $[K_1, \infty)$. In the middle region, we have that

$$\begin{aligned} p^{-1}\rho r H'(z) &= -\rho(a+1) + \tfrac{1}{2}\kappa^2 ba(a+1)\left(\frac{z}{K_1}\right)^{b-1} \\ &< -\rho(a+1) + \tfrac{1}{2}\kappa^2 ba(a+1) \\ &= (a+1)(-\rho + \tfrac{1}{2}\kappa^2 ab) \\ &= 0, \end{aligned}$$

using the product of the roots of the quadratic Q, (A2). Hence H is strictly decreasing through $[0.K_1]$ from a positive value to a negative value, so there is a unique zero. The final assertion of the Proposition follows when we observe that $H(1) > 0$ if and only if $p < p^*$. □

A.3 The budget constraint.

The budget constraint (24) reads

$$w_0 = E \int_0^\infty e^{-\rho t} \tilde{\xi}_t c_t \, dt$$

which we can re-express as

$$\rho w_0 = E\big[\, \tilde{\xi}_T c_T \,\big] = E\Big[\, \tilde{\xi}_T F(\tilde{\xi}_T, \underline{\tilde{\xi}}_T) \,\Big],$$

where we use $\underline{\tilde{\xi}}_t \equiv \min_{0 \le s \le t} \xi_s$, where T is an $\exp(\rho)$ random variable independent of Z, and the form of F can be deduced from (18). We shall therefore need to find the joint law of $(\xi_T, \underline{\tilde{\xi}}_T)$. But if we write $Y_t = \log \tilde{\xi}_t$, then Y is a Brownian motion with drift, and the joint law of[15] $(-\underline{Y}_T, Y_T - \underline{Y}_T)$ is known explicitly; the two random variables are independent exponentials, with parameters a and b respectively. Therefore

$$\begin{aligned} E\Big[\, \tilde{\xi}_T F(\tilde{\xi}_T, \underline{\tilde{\xi}}_T) \,\Big] &= \int_0^\infty \int_0^\infty a e^{-av} b e^{-bs} F(e^{-v+s}, e^{-v}) e^{-v+s} \, dv ds \\ &= \int_0^1 dt \int_1^\infty dy \; ab F(ty, t) t^a y^{-b} \\ &= \frac{2\rho}{\kappa^2} \int_0^1 dt \int_1^\infty dy \; F(ty, t) t^a y^{-b}. \end{aligned}$$

To make the form of the function F explicit, we use the results of Theorem 1. For a general utility, this is about as far as we can go, and the integrals will need to be evaluated numerically, but for the CRRA utility example, we are able to calculate the integrals in closed form with some labour.

[15] Of course, $\underline{Y}_t = \inf_{0 \le s \le t} Y_s$.

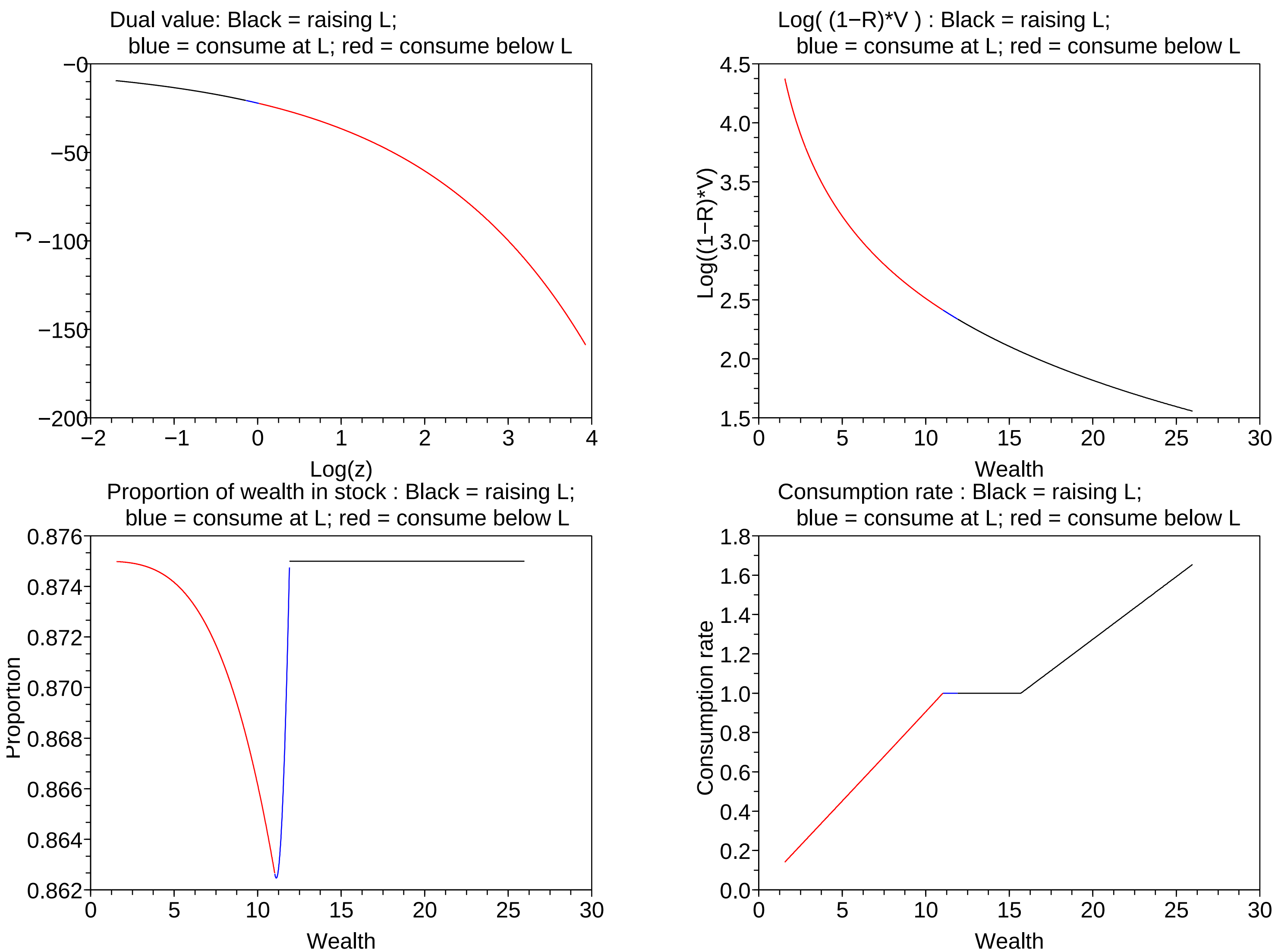


Figure 2: Plots for $K = 1.03$.

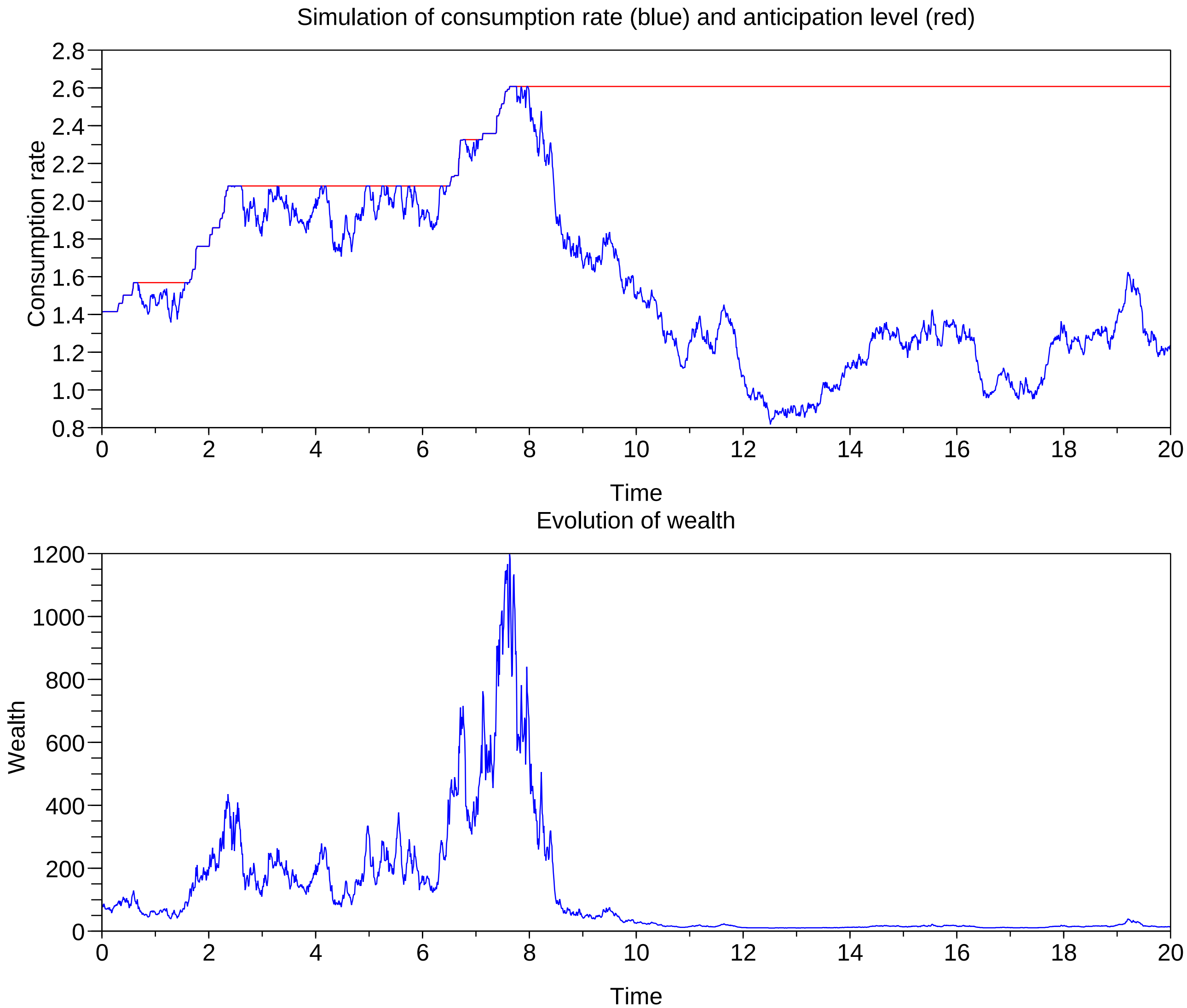


Figure 3: Plots for $K = 1.03$.

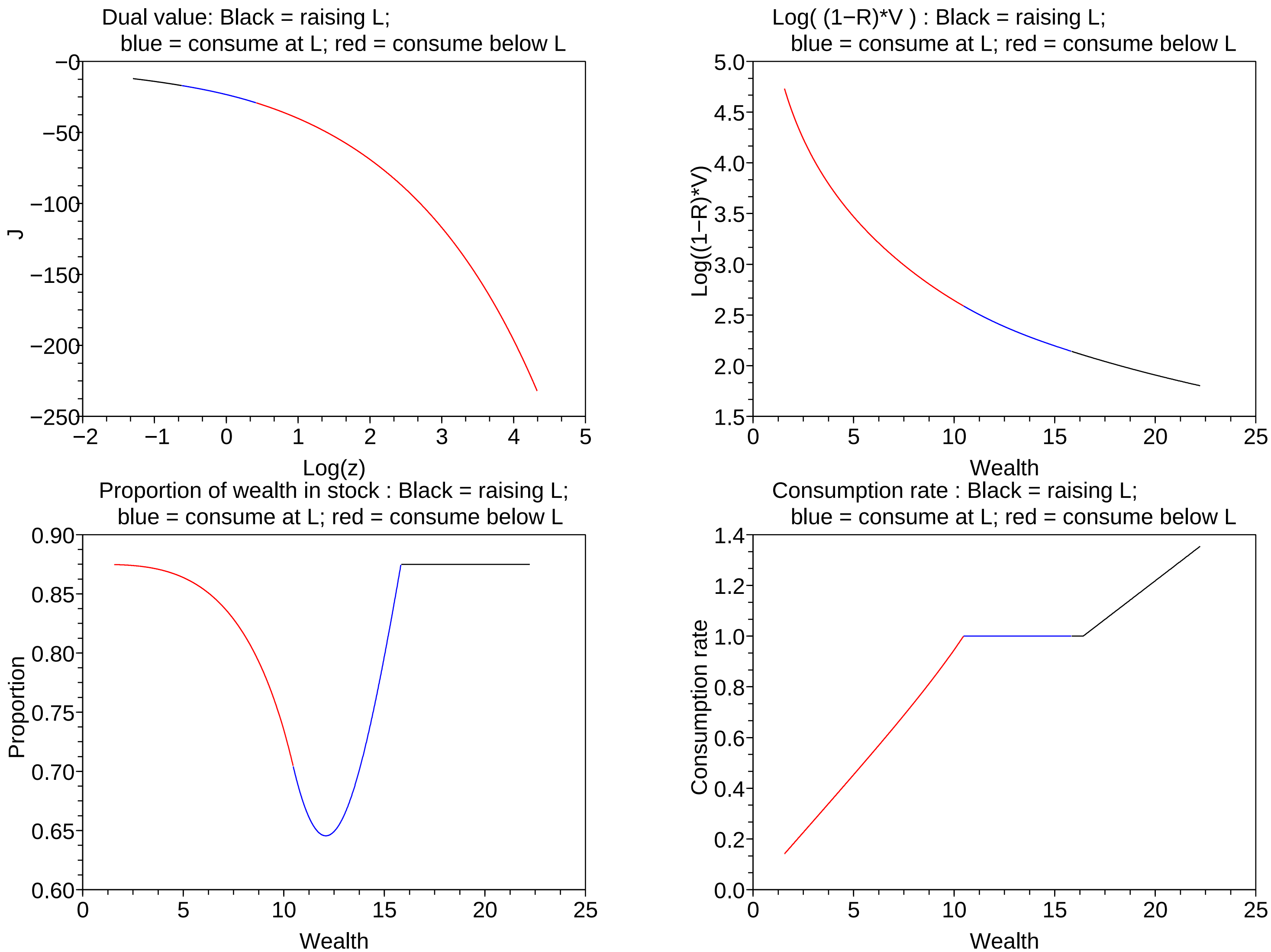


Figure 4: Plots for $K = 2.03$.

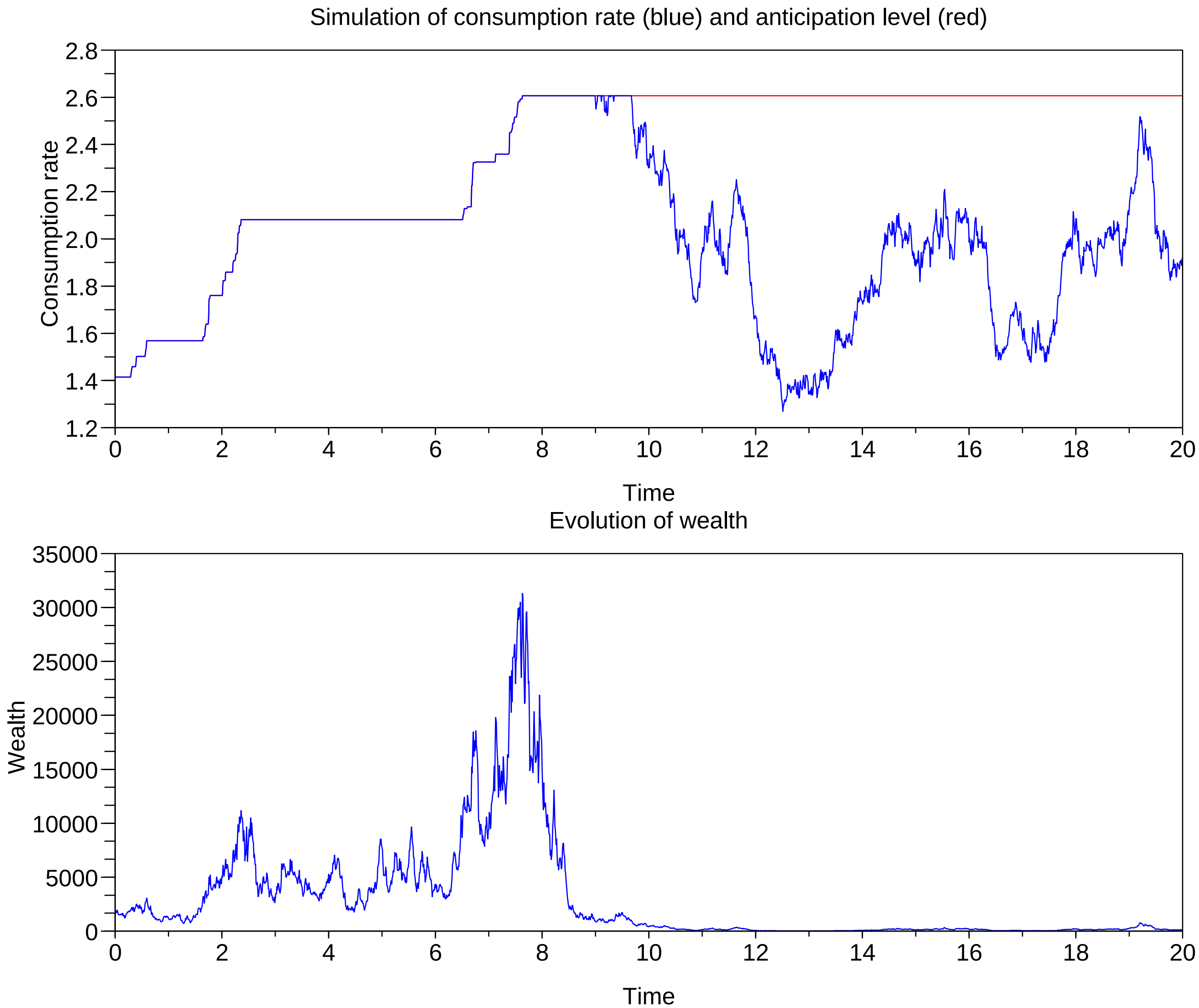


Figure 5: Plots for $K = 2.03$.

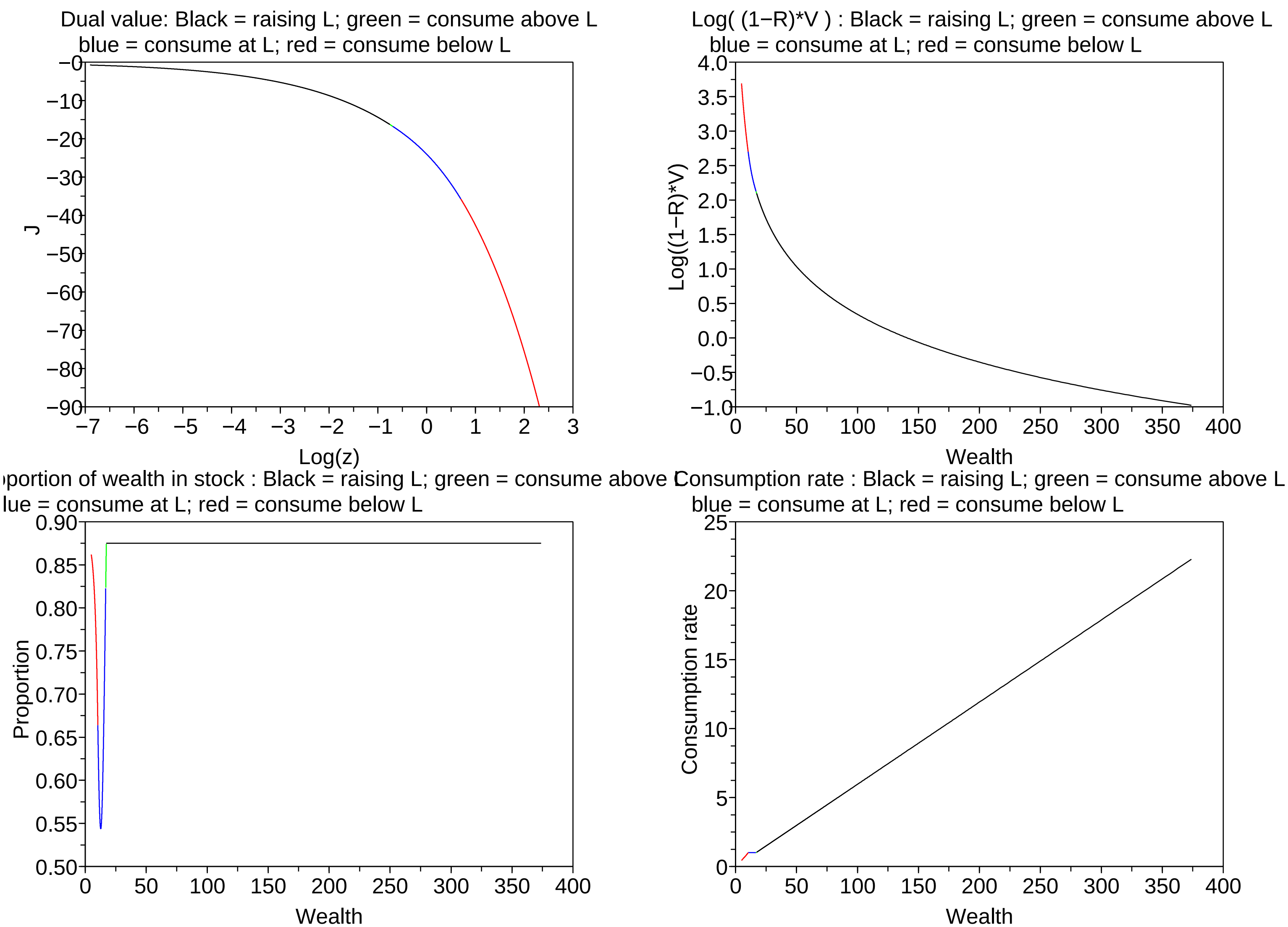


Figure 6: Plots for $K = 3.03$.

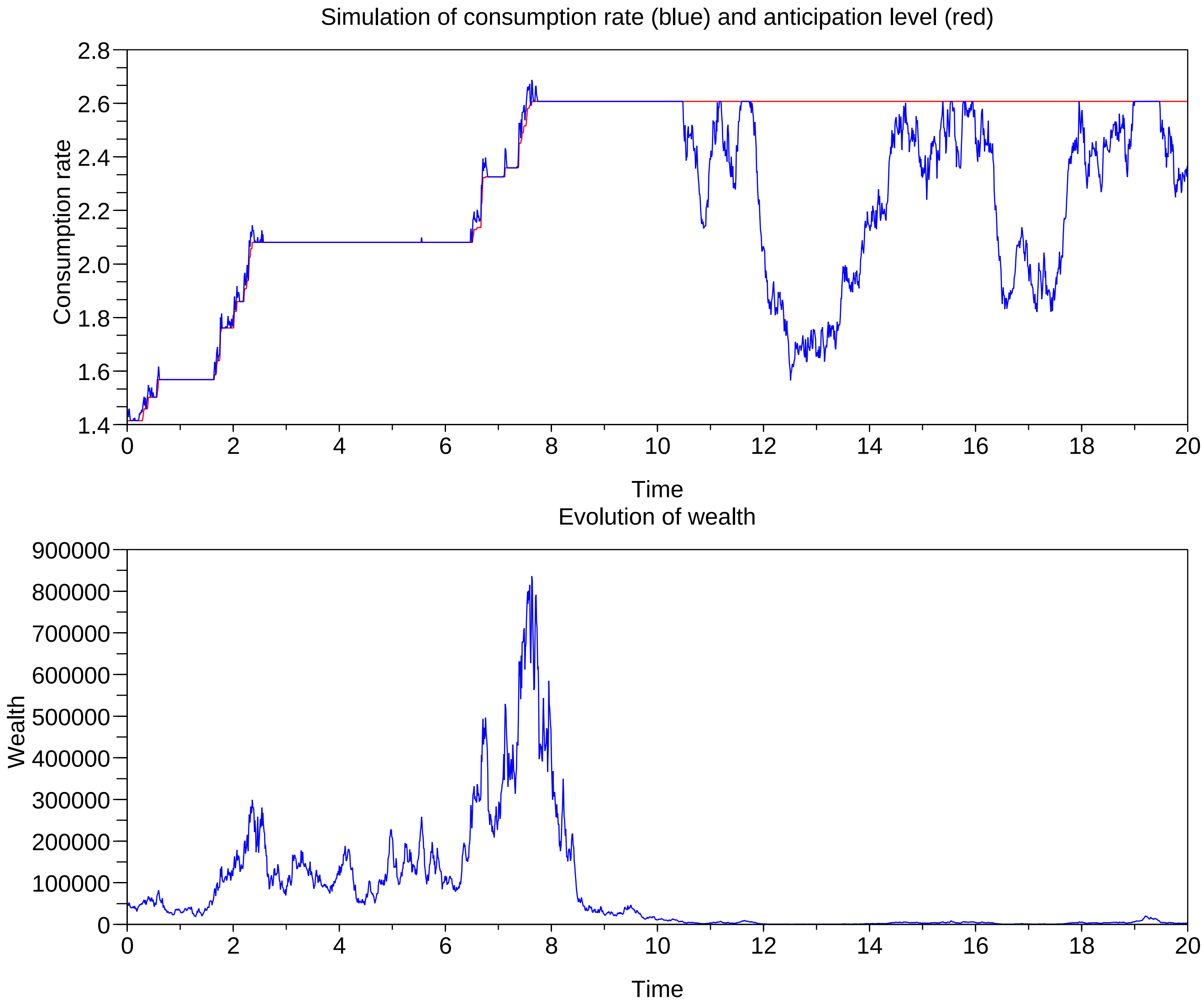


Figure 7: Plots for $K = 3.03$.

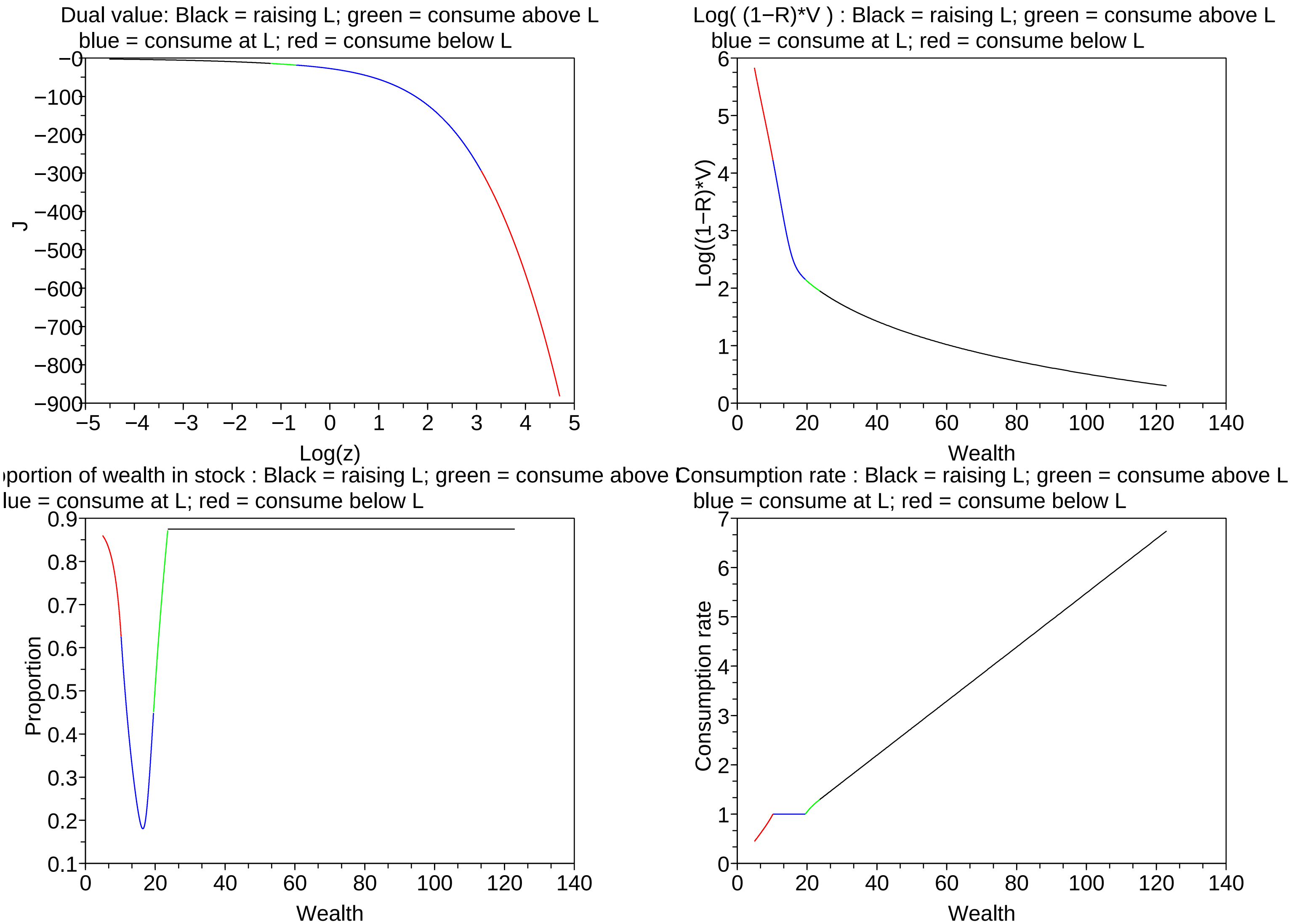


Figure 8: Plots for $K = 43$.

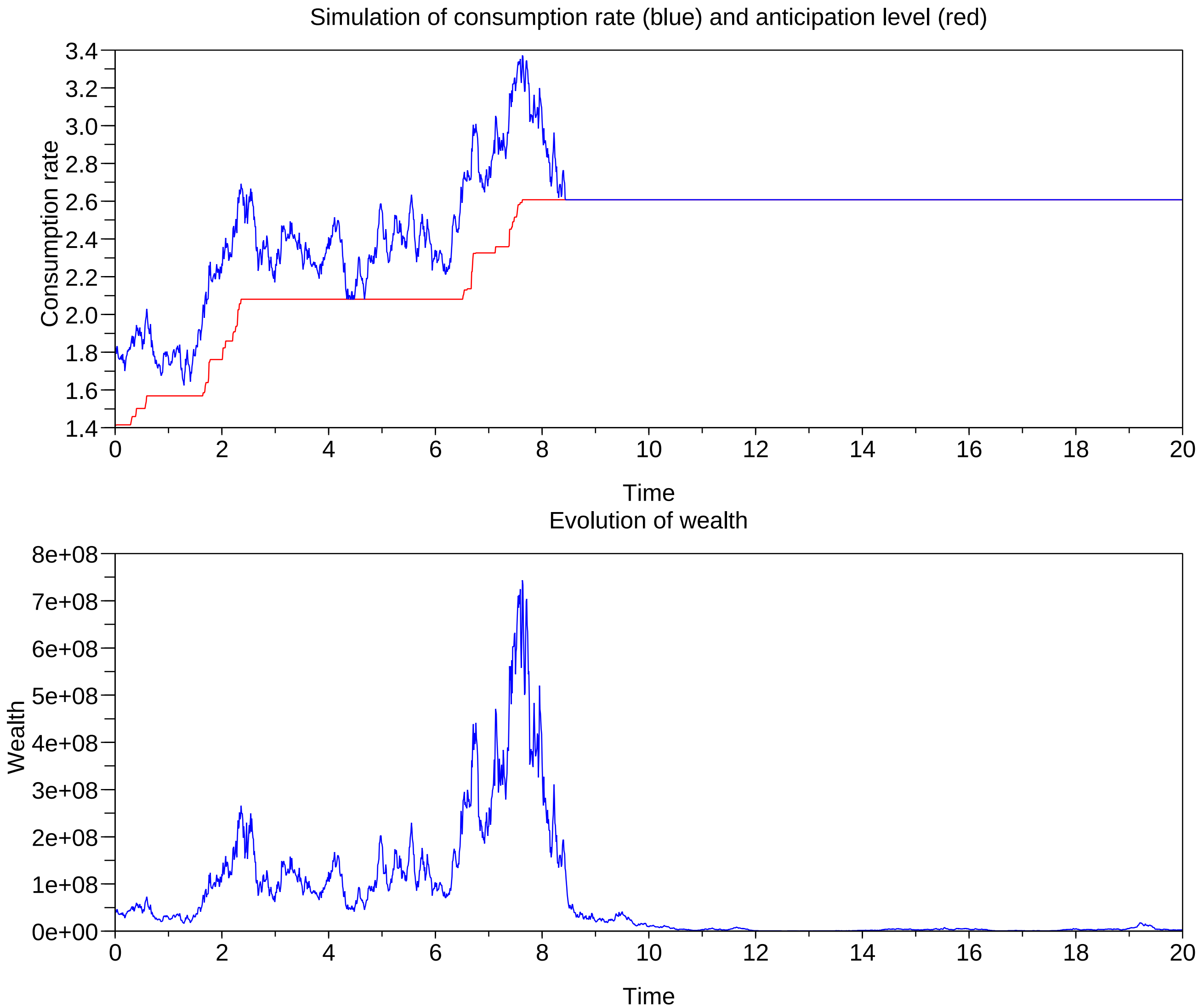


Figure 9: Plots for $K = 43$.