The Fundamental Nature of HARA Utility

Gadi S. Perets*

University of Lyon 1 and ICJ , CNRS

Eran Yashiv†

Tel Aviv University and CEPR

September 2, 2013

Abstract

Many models in Economics *assume* a utility function belonging to the HARA family. This paper shows that HARA utility is fundamental to economic analysis. The HARA functional form is the unique form which satisfies basic economic principles in an optimization context. Using HARA is not just a matter of convenience or tractability but rather an essential restriction.

The paper applies Lie symmetries to the optimality equation of Merton's (1969, 1971) widely-used intertemporal model of consumers-investors. Lie symmetries derive the conditions whereby the optimal solution remains invariant under scale transformations of wealth. The latter arise as the result of growth over time or due to the

*E-mail: gadi@math.univ-lyon1.fr

† *Corresponding author*. E-mail: yashiv@post.tau.ac.il

effects of policy. The symmetries place restrictions on the model, with the key one being the use of HARA utility. We show that this scale invariance of agents' wealth implies linear optimal solutions to consumption and portfolio allocation and linear risk tolerance (and vice versa).

The results have broad implications, as the model studied is a fundamental one in Macroeconomics and Finance. The paper demonstrates the use of Lie symmetries as a powerful tool to deal with economic optimization problems.

Key words: HARA utility, invariance, consumption and portfolio choice, macroeconomics, finance, Lie symmetries.

JEL Codes: D01, D11, D91, E21, G11, C60.

The Fundamental Nature of HARA Utility [1]

1 Introduction

Many models in Economics, and in particular in Macroeconomics and Finance, *assume* a CRRA utility function or some other form belonging to the HARA family. This assumption has proved useful for the tractability of the analysis and has squared well with empirical work. This paper shows that HARA utility is fundamental to economic analysis. The idea is that the HARA functional form is the unique form which satisfies some basic economic principles related to optimization. Hence, using HARA is not just a matter of convenience or tractability, but rather an essential restriction.

More specifically, the analysis establishes an interdependence between the functional form of agents' preferences and the requirement that the optimal solution to a consumption and portfolio choice problem remain invariant under multiplicative wealth transformations. This invariance requirement is fundamental to economic analysis, and the exploration undertaken in this paper shows that it defines what type of utility functions may be used. Relying on such invariance underpins empirical undertakings that aim at estimating stable structural relationships, as advocated by Lucas (1976).

Specifically, the paper shows that by applying Lie symmetries to the optimality equation of Merton's (1969, 1971) widely-used intertemporal model of consumers-investors, the HARA functional form of utility emerges as a basic restriction, rather than just assumed. This means that the HARA form has an economic rationale. We also show that scale invariance of agents' wealth implies linear optimal solutions to the control variables (consumption

[1] *We thank Miles Kimball and Robert Shimer for useful comments. Any errors are our own. We dedicate this paper to the memory of Bill Segal.*

and portfolio allocation) and, vice versa, linear optimal rules imply scale invariance. Indeed, scale invariance determines the relevant linear parameters of optimal behavior. The HARA form itself implies linear risk tolerance. We discuss the connections between linear scale transformations, linear optimal solutions and linear risk tolerance.

Why so, and what is the economic rationale? In order to see that we need to briefly elaborate on some concepts.

First, a *symmetry* is an invariance under transformation. This concept is usually known for the case of the invariance of functions, the homothetic utility or production functions being the most well known special cases. In this paper we use Lie symmetries, which are symmetries of *differential equations*. We derive the symmetries of the optimality equation of the Merton model, which leave the solution of this differential equation invariant.[2]

Second, the *HARA* family of utility functions is a rich one, with absolute or relative risk aversion increasing, decreasing, or constant. It encompasses the special cases of DARA and CARA functions, of the quadratic function, and of the widely-used CRRA functions.[3]

Using these concepts, the underlying rationale of the analysis is the following. Optimal consumption and portfolio choice is subject to variations in scale. Thus, agents have resources of different scale, such as different levels of wealth or income. This could be the result of growth over time or the effects of policy, such as taxation. The model features a differential equation, which expresses the relevant optimality condition (the Bellman equation). Importantly, the latter incorporates the assumption that agents face log-normal asset prices. The solution of this equation defines optimal behavior. Lie symmetries derive the

[2]Other economic implications of Lie symmetries have been discussed and analyzed by a number of authors, and primarily in the pioneering contributions of Sato (1981) and Sato and Ramachandran (1990, 1997). See also the references therein.

[3]See, for example, the discussion in Merton (1971, p.389).

conditions whereby this optimal solution remains invariant under the cited scale transformations of wealth. Doing so, the symmetries impose restrictions on the model, with the key restriction being the use of HARA utility. Note that this is based on an 'if and only if' property: the optimal solution will remain invariant under multiplicative transformations of the agent's wealth *if and only if* HARA utility is used. Hence, if one requires optimal behavior to satisfy certain restrictions – the scale invariance here – then the analysis shows that restrictions need to be placed on the form of the utility function, taking the specific form of HARA utility. This is not to claim that agents need to have HARA utility as their form of preference; what it does say is that economic modelling of optimization which obeys certain economically-relevant invariance requirements, implies these restrictions. This result has broad implications, as the Merton (1969,1971) model is a fundamental one in Macroeconomics and Finance. For example, the Ramsey model and the stochastic growth model, which underlies business cycle modelling, can be thought of as variants of this model in the present context.

The paper proceeds as follows: in Section 2, we briefly introduce the mathematical concept of Lie symmetries of differential equations. In Section 3 we discuss the Merton (1969, 1971) model of consumer-investor choice under uncertainty. In Section 4 we explore the application of Lie symmetries to this model. We derive the main results with respect to HARA utility and discuss them. Section 5 concludes, pointing to further possible uses of this analysis.

2 Lie Symmetries

We briefly present the mathematical concept of Lie symmetries of differential equations,[4] stressing the intuition. For an extensive formal discussion, including applications, see Olver (1993). For an introduction to economic applications see Sato (1981) and Sato and Ramachandran (1990). The aim of the section is to explain the mathematics underlying the analysis below, but the latter may be understood also without the exposition of this section.

2.1 General Concepts of Symmetries and Basic Intuition

To give some general intuition to the concept of symmetries consider first a symmetry of a geometric object. This is a transformation of the space in which it is embedded, which leaves the object invariant. The symmetries of an equilateral triangle, for example, are the rotations in angles $\pi/3,\ 2\pi/3,\ 2\pi$ and the three reflections with respect to the bisectors. The symmetries of a circle centered in the origin are all the possible rotations (angles $0 \leq \alpha \leq 2\pi$) and all the reflections with respect to axes passing through the origin. In each case the symmetries form a group with respect to composition. In the first case this is a discrete (finite) group of six elements (it is actually isomorphic to the group of permutations of three letters) and in the second case this is a continuous group which contains the circle itself as a subgroup. Now consider reversing the order: first, fix the set of symmetries and then see which geometric object 'obeys' this set. In the example of the triangle, if we fix the set of symmetries to be the rotation of angle $\pi/3$ and the reflections around the y-axis (and indeed all the possible combinations of these symmetries, hence the group generated by the two symmetries), we obtain that the only geometric object preserved by these two symmetries

[4] Named after the Norwegian mathematician Sophus Lie, whose work dates back to the late 19th century.

is an equilateral triangle. This establishes a dual way to 'see' a triangle, i.e., through its symmetries. The same could be done with the circle, the only difference being that the set of symmetries preserving the circle consists of all possible rotations (with angles $0 \leq \alpha \leq 2\pi$) and all possible reflections with respect to axes passing through the origin. This difference, however, is a significant one as it introduces continuity: the group of symmetries being a continuous group, we are now permitted to use concepts of continuous mathematics in order to understand the interplay between the geometric object and the set of symmetries preserving it.

A similar continuous approach may be used to analyze a system of differential equations, *which is what is to be done in this paper.* We can view a system of partial differential equations as a description of a geometric object which is the space of solutions of the system. The symmetries of the differential equations are thus the transformations which leave the space of solutions invariant. Determining the group of symmetries of the space of solutions may give valuable insights with respect to the solution itself. As shown by the example below, for the p.d.e $\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2}$, the heat equation in Physics, the group of symmetries of the equation not only gives insight to the problem in question but actually provides a way to get to the solution itself.

The power of this theory lies in the notion of infinitesimal invariance: one can replace complicated, possibly highly non-linear conditions for invariance of a system by equivalent linear conditions of infinitesimal invariance. Infinitesimal symmetries are elements of the tangent space to symmetries of the system. To employ familiar concepts, it is analogous to the use of derivatives of a function at a point to approximate the function in the neighborhood of this point. Likewise the infinitesimal symmetries are "derivatives" of the actual symmetries. The way to go back from the former to the latter is through an exponentiation procedure. To use familiar terminology again, the latter is analogous to the use of a Taylor series.

A crucial point is that if one is looking for smooth symmetries and the equations in question satisfy some non-degeneracy conditions (as is the case analyzed in this paper) then *all* the smooth symmetries of the equation system are derived through the infinitesimal symmetries. We stress this point as the symmetries derived below express diverse aspects of the consumer-investor optimization problem. The afore-cited property assures us of extracting all the possible symmetries.

2.2 Invariance of Differential Equations

The Lie symmetries of differential equations are the transformations which leave the space of solutions invariant. We begin by explaining the concept of invariance of differential equations, culminating by the derivation of the prolongation equation, which is key in deriving the Lie symmetries of a differential equations system.

Consider the differential equation:

$$L(t, x, y, p) = 0 \tag{1}$$

where $x = x(t), y = y(x), p = \frac{dy}{dx}$ and t is time.

The transformation:

$$\begin{aligned} x' &= \phi(x, y, t) \\ y' &= \psi(x, y, t) \end{aligned}$$

implies the transformation of the derivative $p = \frac{dy}{dx}$ to:

$$p' = \frac{dy'}{dx'} = \frac{\frac{\partial\psi}{\partial x}dx + \frac{\partial\psi}{\partial y}dy}{\frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy} = \frac{\frac{\partial\psi}{\partial x} + \frac{\partial\psi}{\partial y}p}{\frac{\partial\phi}{\partial x} + \frac{\partial\phi}{\partial y}p} \tag{2}$$

The differential equation (1) will be invariant under the transformation $x \to x'$ and $y \to y'$ (i.e., one integral curve is mapped to another) if and only if it is invariant under:

$$\begin{aligned} x' &= \phi(x, y, t) \\ y' &= \psi(x, y, t) \\ p' &= \chi(x, y, p, t) \end{aligned}$$

The condition for transformation (3) to leave the differential equation (1) invariant is:

$$H'L \equiv \xi\frac{\partial L}{\partial x} + \eta\frac{\partial L}{\partial y} + \eta'\frac{\partial L}{\partial p} = 0 \tag{3}$$

where:

$$\begin{aligned} H &= (\frac{\partial\phi}{\partial t})_0\frac{\partial}{\partial x} + (\frac{\partial\psi}{\partial t})_0\frac{\partial}{\partial y} \\ &= \xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} \\ \xi &\equiv (\frac{\partial\phi}{\partial t})_0 \quad \eta \equiv (\frac{\partial\psi}{\partial t})_0 \\ \eta' &\equiv \frac{\partial\eta}{\partial x} + (\frac{\partial\eta}{\partial y} - \frac{\partial\xi}{\partial x})p - \frac{\partial\xi}{\partial y}p^2 \end{aligned}$$

and the subscript 0 denotes the derivative at $t = 0$; the notation $\frac{\partial}{\partial x}$ is used for a directional derivative i.e., the derivative of the function in the direction of the relevant coordinate axis, assuming space is coordinated. For this and other technical concepts, see Chapter 1 in Sato and Ramachandran (1990).

Below we use an equation like (3) to derive the symmetries of the optimality condition of the Merton (1969,1971) model. To see the intuition underlying equation (3) consider the invariance of a function (a generalization of homotheticity) rather than that of a differential equation: a function $f(x, y)$ is invariant under a transformation $x \rightarrow x'$ and $y \rightarrow y'$ if $f(x, y) = f(x', y')$. Using a Taylor series and infinitesimal transformations we can write:

$$f(x', y') = f(x, y) = f(x, y) - tHf + \frac{t^2}{2}H^2 f + ...$$

It is evident that the necessary and sufficient condition for invariance in this case is:

$$Hf = 0 \tag{4}$$

Equation (3) is the analog of equation (4) for the case of a differential equation. It is called the *prolongation equation* and it is linear in ξ and η. Finding the solution to it gives the infinitesimal symmetries from which the symmetries of the differential equation itself may be deduced.

2.3 An Example of Lie Symmetries: the Heat Equation

As a concrete example of Lie symmetries, making use of the above methodology, we shall now consider a key application in Physics, pertaining to the equation for the conduction of heat in a one-dimensional rod. This is *not* to say that the economic analysis given below is akin to a problem in Physics, but rather to give as an example a key application of the mathematical technique. This example is helpful but not essential to understanding the economic analysis below.

As noted above, the power of this theory lies in the notion of infinitesimal invariance: one can replace complicated, possibly highly non-linear conditions for invariance of a system by equivalent linear conditions of infinitesimal invariance and is analogous to the use of derivatives of a function at a point to approximate the function in the neighborhood of this point. Likewise, the infinitesimal symmetries are "derivatives" of the actual symmetries and the way to go back from the former to the latter is through an exponentiation procedure.

The differential equation. Denote $u(x,t)$ as the temperature of a rod in the x coordinate at time t. The heat equation is given by the following partial differential equation:[5]

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2} \tag{5}$$

Note that there is no straightforward solution to equation (5).

The infinitesimal generators. The Lie symmetries of this equation are transformations which leave the solution set of equation (5) invariant. We find them by calculating their infinitesimal generators, which are vector fields on the manifold composed of all the invariance transformations. Finding these generators is relatively easy, as it is more of an algebraic calculation, while finding the invariance transformations directly amounts more to an analytic calculation. After finding the infinitesimal generators, we "exponentiate" them to get the actual invariant transformations.

A general infinitesimal generator is of the form:

$$v = \xi(x,t,u)\frac{\partial}{\partial x} + \tau(x,t,u)\frac{\partial}{\partial t} + \phi(x,t,u)\frac{\partial}{\partial u} \tag{6}$$

[5]The thermal diffusivity having been normalized to unity.

We determine all the possible functions ξ, τ, ϕ through the prolongation equation (see equation (3) above), which puts together all the possible constraints on the functions ξ, τ, ϕ.

Note that if a group G is a Lie group[6], then there are certain vector fields on G, characterized by their invariance under group multiplication. These invariant vector fields form a finite-dimensional vector space called the Lie algebra of G, the elements of which are the infinitesimal generators of G. Here we get that all infinitesimal generators form a Lie algebra, i.e., a Lie product of two generators gives an infinitesimal generator. The prolongation formula alluded to above yields, after some calculations, the basis of this Lie algebra in the current case as follows:

[6] Olver (1993, pages 14-15) offers these definitions: a group is "a set G together with a group operation, usually called multiplication, such that for any two elements g and h of G the product $g \bullet h$ is again an element of G...

An r-parameter Lie group is a group G which also carries the structure of an r-dimensional smooth manifold in such a way that both the group operation

$m : G \times G \to G, \ \ m(g,h) = g \bullet h, \ \ g,h \in G$

and the inversion

$i : G \to G, \ \ i(g) = g^{-1}, \ \ g \in G$

are smooth maps between manifolds."

$$\begin{aligned}
\nu_1 &= \delta_x \\
\nu_2 &= \delta_t \\
\nu_3 &= u\delta_u \\
\nu_4 &= x\delta_x + 2t\delta_t \\
\nu_5 &= 2t\delta_x - xu\delta_u \\
\nu_6 &= 4tx\delta_x + 4t^2\delta_t - (x^2+2t)u\delta_u \\
\nu_\alpha &= \alpha(x,t)\delta_u
\end{aligned}$$

Exponentiation and the symmetries. Following exponentiation of the infinitesimal generators, we get that the group of invariance transformations is generated by the following transformations.

If:

$$u = f(x,t) \tag{7}$$

is a solution to the heat equation (5), then so are:

$$
\begin{aligned}
u^{(1)} &= f(x-\epsilon, t) \\
u^{(2)} &= f(x, t-\epsilon) \\
u^{(3)} &= e^{\epsilon} f(x,t) \\
u^{(4)} &= f(e^{-\epsilon}x, e^{-2\epsilon}t) \\
u^{(5)} &= e^{-\epsilon x+\epsilon^2 t} f(x-2\epsilon t, t) \\
u^{(6)} &= \frac{1}{\sqrt{1+\epsilon t}} exp\left\{\frac{-\epsilon x^2}{1+4\epsilon t}\right\} f(\frac{x}{1+4\epsilon t}, \frac{t}{1+4\epsilon t}) \\
u^{(\alpha)} &= f(x,t) + \epsilon\alpha(x,t)
\end{aligned} \tag{8}
$$

Henceforth we shall call the above (equations (8)) – symmetry groups. The symmetry groups G_3 and G_α reflect the linearity of the heat equation, i.e., we can add up solutions and multiply them by constants. The groups G_1 and G_2 demonstrate the time and space invariance to the equation, reflection of the fact that the heat equation has constant coefficients. The scaling symmetry is G_4, while G_5 represents a Galilean transformation of the moving coordinate frame.

The group G_6 is a genuine group of transformations. Its existence is far from obvious from basic physical principles and it has the following far-reaching implication:

The constant function $u = c$ is a trivial solution. But if we apply the sixth symmetry above to it, then we conclude that the function

$$u = \frac{c}{\sqrt{1+4\epsilon t}} exp^{\frac{-x^2}{4t}}$$

is also a solution, which is a fundamental non-trivial solution of the heat equation. This exemplifies the idea that Lie symmetries may yield solutions even when differential equations cannot be solved, as is the case here.

Hence, through the analysis of the invariance transformations of equation (5) we are able to completely solve it. This shows that the analysis of the space of invariance transformations can lead to the understanding of deep implications of the equation itself. While it is not always possible to integrate the solution completely through such analysis, in many case it leads to powerful insights about it.

The above procedure is the one we undertake below for the Hamilton-Jacobi-Bellman optimality equation of the Merton (1969, 1971) model, to which we turn now.

3 Merton's Model of Optimal Consumption and Portfolio Selection

In this section we briefly present the main ingredients of the consumer/investor optimization problem under uncertainty as initially formulated and solved by Merton (1969, 1971)[7]. We choose this model as it is a fundamental model of consumer/investor choice and is akin to other prevalent models, such as the Ramsey model or the stochastic growth model The essential problem is that of an individual who chooses an optimal path of consumption and portfolio allocation. The agent begins with an initial endowment and during his/her lifetime consumes and invests in a portfolio of assets (risky and riskless). The goal is to maximize the expected utility of consumption over the planning horizon and a "bequest" function defined in terms of terminal wealth.

Formally the problem may be formulated in continuous time, using Merton's notation,

[7]For a discussion of developments since the initial exposition of these papers see Merton (1990, chapter 6) and Duffie (2003). For some discussion of application of Lie symmetries to portfolio problems of this kind, but not the one suggested here, see Boyd (1990).

as follows: denote consumption by C, financial wealth by W, time by t (running from 0 to T), utility by U, and the bequest by B. There are two assets used for investment,[8] one of which is riskless, yielding an instantaneous rate of return r. The other asset is risky, its price P generated by an Ito process as follows:

$$\frac{dP}{P} = \alpha(P,t)dt + \sigma(P,t)dz \tag{9}$$

where α is the instantaneous conditional expected percentage change in price per unit time and σ^2 is the instantaneous conditional variance per unit time.

The consumer seeks to determine optimal consumption and portfolio shares according to the following:

$$\max_{(C,w)} E_0 \left\{ \int_0^T U[C(t),t]dt + B[W(T),T] \right\} \tag{10}$$

subject to

$$dW = w(\alpha - r)Wdt + (rW - C)dt + W\sigma dz \tag{11}$$

$$W(0) = W_0 \tag{12}$$

[8]The problem can be solved with n risky assets and one riskless asset. As in Merton (1971) and for the sake of expositional simplicity, we restrict attention to two assets. Our results apply to the more general case as well.

where w is the portfolio share invested in the risky asset. All that needs to be assumed about preferences is that U is a strictly concave function in C and that B is concave in W.

Merton (1969, 1971) applied stochastic dynamic programming to solve the above problem. In what follows we repeat the main equations; see Sections 4-6 of Merton (1971) for a full derivation.

Define:

(i) An "indirect" utility function:

$$J(W, P, t) \equiv \max_{(C,w)} E_t \left\{ \int_t^T U(C, s)ds + B[W(T), T] \right\} \tag{13}$$

where E_t is the conditional expectation operator, conditional on $W(t) = W$ and $P(t) = P$.

(ii) The inverse marginal utility function:

$$G \equiv [\partial U/\partial C]^{-1} \equiv {U_C}^{-1}(C) \tag{14}$$

The following notation will be used for partial derivatives: $U_C \equiv \partial U/\partial C$, $J_W \equiv \partial J/\partial W$, $J_{WW} \equiv \partial^2 J/\partial W^2$, and $J_t \equiv \partial J/\partial t$.

A sufficient condition for a unique interior maximum is that $J_{WW} < 0$ i.e., that J be strictly concave in W.

Merton assumes "geometric Brownian motion" holds for the risky asset price, so α and σ are constants and prices are distributed log-normal. In this case J is independent of P, i.e., $J = J(W, t)$.

Time preference is introduced by incorporating a subjective discount rate ρ into the utility function:

$$U(C,t) = \exp(-\rho t)\widetilde{U}(C,t) \tag{15}$$

The optimal conditions are given by:

$$\exp(-\rho t)\widetilde{U}_C(C^*,t) = J_W \tag{16}$$

$$(\alpha - r)WJ_W + J_{WW}w^*W^2\sigma^2 = 0 \tag{17}$$

where C^*, w^* are the optimal values.

Combining these conditions results in the so-called Hamilton-Jacobi-Bellman (HJB) equation, which is a partial differential equation for J, one obtains:

$$U(G,t) + J_t + J_W\,(rW - G) - \frac{J_W^2}{J_{WW}}\frac{(\alpha - r)^2}{2\sigma^2} = 0 \tag{18}$$

subject to the boundary condition $J(W,T) = B(W,T)$. Merton (1971) solved the equation by restricting preferences, *assuming* that the utility function for the individual is a member of the Hyperbolic Absolute Risk Aversion (HARA) family of utility functions. The optimal C^* and w^* are then solved for as functions of J_W and J_{WW}, the riskless rate r, wealth W, and the parameters of the model (α and σ^2 of the price equation and the HARA parameters).

4 The Symmetries of the Consumer-Investor Optimality Equation and Their Economic Interpretation

We now derive the symmetries of the HJB equation (18). Two issues should be emphasized: (i) the symmetries are derived with no assumption on the functional form of the utility function except its concavity in C, a necessary condition for maximization; (ii) the optimal solution depends on the derivatives of the indirect utility function J, which, in turn, depends on wealth W and time t. The idea is to derive transformations of t and W that would leave the optimality *equation* invariant. These transformations do not require imposing any restrictions on the end points, i.e., transversality conditions, of the type usually needed to obtain a unique solution to optimal control problems.[9]

In economic terms, this means that if wealth varies, say because of taxation or because of intertemporal growth, the optimal solution remains invariant. The underlying interest in the invariance of the optimality equations is that we would like to have invariance of the structure of the solution across different levels of wealth.

In what follows we present the derivation of the symmetries (4.1) and their economic interpretation (4.2).

4.1 Derivation of the Symmetries

4.1.1 The Infinitesimal Generators of the Symmetries

In order to calculate the symmetries of the HJB equation (18), we first calculate the infinitesimal generators of the symmetries, and then exponentiate these infinitesimal generators

[9] The symmetries, however, do not restrict the optimal solution to be unique.

to get the symmetries themselves. An infinitesimal generator ν of the HJB equation has the following form, as in equation (6) above:

$$\nu = \xi(W,t,J)\frac{\partial}{\partial W} + \tau(W,t,J)\frac{\partial}{\partial t} + \phi(W,t,J)\frac{\partial}{\partial J} \tag{19}$$

Here ξ, τ, ϕ are functions of the variables W, t, J. The function J, as well as its partial derivatives, become variables in this method of derivation of the symmetries. In order to determine explicitly the functions ξ, τ, ϕ we prolongate the infinitesimal generator ν according to the prolongation formula of Olver (1993, page 110) and the equations thereby obtained provide the set of constraints satisfied by the functions ξ, τ, ϕ (see details in Olver (1993, pages 110-114), whose notation we use throughout).

The prolongation equation (see equation (3) above) applied to ν yields:

$$\begin{aligned} & \left[r\xi J_W - \rho\tau e^{-\rho t}U(G(e^{\rho t}J_W)) + (rW - G(e^{\rho t}J_W))\phi^W + \phi^t\right] J_{WW}^2 \\ & +2A\phi^W J_W J_{WW} - A\phi^{WW} J_W^2 \\ = \; & 0 \end{aligned}$$

where $\phi^W, \phi^t, \phi^{WW}$ are given by:

$$\phi^W = \phi_W + (\phi_J - \xi_W)J_W - \tau_W J_t - \xi_J J_W^2 - \tau_J J_W J_t$$

$$\phi^t = \phi_t - \xi_t J_W + (\phi_J - \tau_t)J_t - \xi_J J_W J_t - \tau_J J_t^2$$

$$\begin{aligned} \phi^{WW} = \; & \phi_{WW} + (2\phi_{WJ} - \xi_{WW})J_W - \tau_{WW}J_t + (\phi_{JJ} - 2\xi_{WJ})J_W^2 \\ & -2\tau_{WJ}J_W J_t - \xi_{JJ}J_W^3 - \tau_{JJ}J_W^2 J_t + (\phi_J - 2\xi_W)J_{WW} \\ & -2\tau_W J_{Wt} - 3\xi_J J_W J_{WW} - \tau_J J_t J_{WW} - 2\tau_J J_W J_{Wt} \end{aligned}$$

Plugging in these in the above equation we get:

$$
\begin{aligned}
& r\xi J_W - \rho\tau e^{-\rho t} U(G(e^{\rho t} J_W)) \\
& +(rW - G(e^{\rho t} J_W))(\phi_W + (\phi_J - \xi_W) J_W - \tau_W J_t - \xi_J J_W^2 - \tau_J J_W J_t) \\
& +(\phi_t - \xi_t J_W + (\phi_J - \tau_t) J_t - \xi_J J_W J_t - \tau_J J_t^2) J_{WW}^2 \\
& +2A(\phi_W + (\phi_J - \xi_W) J_W - \tau_W J_t - \xi_J J_W^2 - \tau_J J_W J_t) J_W J_{WW} \\
& -A \begin{pmatrix} \phi_{WW} \\ +(2\phi_{WJ} - \xi_{WW}) J_W \\ -\tau_{WW} J_t \\ +(\phi_{JJ} - 2\xi_{WJ}) J_W^2 \\ -2\tau_{WJ} J_W J_t \\ -\xi_{JJ} J_W^3 \\ -\tau_{JJ} J_W^2 J_t \\ +(\phi_J - 2\xi_W) J_{WW} \\ -2\tau_W J_{Wt} \\ -3\xi_J J_W J_{WW} \\ -\tau_J J_t J_{WW} \\ -2\tau_J J_W J_{Wt} \end{pmatrix} J_W^2 \\
= \;& 0
\end{aligned}
\tag{20}
$$

Note that the variables J_W, J_{WW}, J_{Wt}, J_t are algebraically independent. This implies that the coefficients of the different monomials in those variables are equal to zero. We therefore proceed as follows.

(i) We first look at the different monomials in the above equation in which J_{WW} does not appear. Equating the coefficients of these monomials to 0 implies that:

$$\begin{aligned} \tau_J &= \tau_W = 0 \\ \xi_{JJ} &= 0 \\ \phi_{WW} &= 0 \\ \phi_{JJ} - 2\xi_{JW} &= 0 \\ 2\phi_{WJ} - \xi_{WW} &= 0 \end{aligned} \tag{21}$$

(ii) Next we look at monomials in which J_{WW} appears in degree one. This gives (noting that $r \neq \alpha$ implies that $A \neq 0$) the following equation (in which we gathered only those monomials with their coefficients):

$$2(\phi_W + (\phi_J - \xi_W)J_W - \xi_J J_W^2)J_W J_{WW} + 3\xi_J J_W^3 J_{WW} - (\phi_J - 2\xi_W)J_{WW}J_W^2 = 0$$

From this we deduce that

$$\begin{aligned} \phi_W &= 0 \\ \xi_J &= 0 \\ \phi_J &= 0 \end{aligned} \tag{22}$$

(iii) Now we look at the monomials containing J_{WW}^2 which give the following equation:

$$r\xi J_W - \rho\tau e^{-\rho t}U(G(e^{\rho t}J_W)) - (rW - G(e^{\rho t}J_W))\xi_W J_W + \phi_t - \xi_t J_W - \tau_t J_t = 0 \tag{23}$$

From which we deduce that :

$$\begin{aligned} \phi_t &= 0 \\ \tau_t &= 0 \end{aligned} \tag{24}$$

From all the constraints above on the functions ξ, τ, ϕ we gather so far that $\tau = Constant$ and $\phi = Constant$ and we are left with the following equation for the ξ function:

$$e^{\rho t}(r\xi - \xi_t - rW\xi_W)J_W + \xi_W G(e^{\rho t}J_W)e^{\rho t}J_W - \rho\tau U(G(e^{\rho t}J_W)) = 0 \tag{25}$$

From this we deduce that $\xi_W = 0$ unless the following functional equation is satisfied

$$G(e^{\rho t}J_W)e^{\rho t}J_W - \varepsilon U(G(e^{\rho t}J_W)) = 0 \tag{26}$$

in which ε is a constant scalar. The last statement is of great importance in the current context, as will be shown below.

We end up with the following constraints for the infinitesimal generators:

$$\phi_t = 0 \tag{27}$$

$$\rho\tau = \phi \tag{28}$$

$$\xi_W = 0 \tag{29}$$

The last equation holds true unless equation (26) is satisfied.

4.1.2 The Symmetries

The constraints, which we have derived, on the functions ξ, τ, ϕ and their derivatives completely determine the infinitesimal symmetries, which are given by:

Symmetry 1 $\phi_t = 0$

If $J(W,t)$ is a solution to the HJB equation, then so is $J(W,t) + k$ for any $k \in R$.

Symmetry 2 $\rho\tau = \phi$

If $J(W,t)$ is a solution of the HJB equation, so is $e^{-\rho\tau}J(W, t+\tau)$ for any $\tau \in R$.

Symmetry 3 $\xi_W = 0$

If $J(W,t)$ is a solution of the HJB equation, so is $J(W + ke^{rt}, t)$ for any $k \in R$.

For a general specification of the utility function, the HJB equation of the model admits only the above three symmetries. However, from the constraints above we also get that in case that the utility function satisfies the functional equation (26), and only in that case, there is an extra symmetry for the equation.

We now consider the implications of equation (26). For this we need first the following.

Lemma 1 *The functional equation*

$$G(x)x - \varepsilon U(G(x)) = 0 \tag{30}$$

where $G = (U')^{-1}$ *, is satisfied by a utility function* U *iff* U *is of the HARA form.*

Proof. Upon plugging in the equation a utility $U(x)$ of the HARA form we see the functional equation above is satisfied. Going the other way, after differentiating the equation with respect to x, we get the ordinary differential equation:

$$G'x + G - \varepsilon(xG') = 0$$

The solutions of this equation form the HARA class of utility functions. Then we take the inverse function to get U' and after integration we get that U is of the HARA form. ■

When the functional equation (26) holds true and Lemma 1 is relevant, the exponentiation of the infinitesimal generators yields a fourth symmetry as follows.

Symmetry 4

If $J(W,t)$ is a solution, then so is $e^{k\gamma}J(e^{-k}\{W+\frac{(1-\gamma)\eta}{\beta r}\}-\eta\frac{(1-\gamma)}{\beta r},t)$ for any $k \in R$.

In words this fourth symmetry says that if $J(W,t)$ is a solution then a linear function of J is also a solution. Calculation of the solutions to the functional equation in Lemma 1 shows that only utility functions of the HARA class satisfy it. The HARA function is expressed as follows:

$$\widetilde{U}(C) = \frac{1-\gamma}{\gamma}\left(\frac{\beta C}{1-\gamma}+\eta\right)^{\gamma} \tag{31}$$

The special case of the CRRA function $\frac{C^\gamma}{\gamma}$ has $\beta = 1, \eta = 0$.

4.2 Economic Interpretation

There are four symmetries all together. The first three formulate "classical" principles of utility theory; the fourth places restrictions on the utility function, and it is the main point of interest here.

Symmetry 1

If $J(W,t)$ is a solution to the HJB equation, then so is $J(W,t)+k$ for any $k \in R$.

This symmetry represents a formulation of the idea that utility is ordinal and not cardinal.

Symmetry 2

If $J(W,t)$ is a solution of the HJB equation, so is $e^{-\rho\tau}J(W,t+\tau)$ for any $\tau \in R$.

This symmetry expresses the property that displacement in calendar time does not change the optimal solution.

Symmetry 3

If $J(W,t)$ is a solution for the HJB equation, then so is $J(W+ke^{rt},t)$ for any $k \in R$.

This symmetry expresses a property with respect to W that is similar to the property of Symmetry 2 with respect to t : if the solution is optimal for W then it is also optimal for an additive re-scaling of W; the term e^{rt} keeps the additive k constant in present value terms.

As noted, for a general specification of the utility function, the HJB equation of the model admits only the above three symmetries.

Symmetry 4

If $J(W,t)$ is a solution, then so is $e^{k\gamma}J(e^{-k}\{W+\frac{(1-\gamma)\eta}{\beta r}\}-\eta\frac{(1-\gamma)}{\beta r},t)$ for any $k \in R$.

As noted, this holds only if equation (26) is satisfied.

This fourth symmetry is the key point of this paper. It has the following major implications:

(i) Because k is completely arbitrary any multiplicative transformations of W, i.e., $e^{-k}W$, apply. Such transformations are the most natural to consider when thinking of wealth growth or policy effects.[10]Thus, this symmetry states the following: the optimum, expressed by the J function, i.e., maximum expected life-time utility, will remain invariant under

[10]For a recent empirical discussion of wealth changes and their effects on portfolio allocations with special reference to implications for risk aversion, and therefore for the functional form of the utility function, see Brunnermeier and Nagel (2008).

multiplicative transformations of wealth W *if and only if* HARA utility is used. Hence the HARA form is determined by the symmetry. Note well that HARA utility is implied by this symmetry, not assumed a-priori. This does *not* imply, though, that there is a unique such J but it does express a property of any J which solves the HJB equation.[11]

The idea, then, is that there is an interdependence between the functional form of preferences (the form of the utility function) and the requirement that the optimal solution will remain invariant under multiplicative wealth transformations. This interdependence takes specific form in Symmetry 4.

The iff property means that if we demand scaling invariance of wealth then the utility function has to be HARA and if the utility function is HARA then we have wealth scaling invariance. This kind of invariance underpins empirical undertakings that aim at estimating stable structural relationships, as forcefully advocated by Lucas (1976).

(ii) In Merton (1969, p. 391) the following theorem is presented and proved:

THEOREM III. Given the model specified... then $C^* = aW + b$ and $w^*W = gW + h$ where a, b, g, and h are, at most, functions of time if and only if $U(C,t) \subset \ HARA(C)$.

The result we obtain above can also be stated as follows:

Theorem 2 *Given the model specified in this section, then symmetry 4 (the scaling symmetry) is satisfied if and only if* $U(C,t) \subset \ HARA(C)$.

Combining the last theorem with Merton's theorem above, we get:

Corollary 3 *Given the model specified in this section, then* $C^* = aW+b$ *and* $w^*W = gW+h$

[11]If, together with the multiplicative transformation, there is also an additive transformation of W, expressed by the term $e^{-k}\left(\frac{(1-\gamma)\eta}{\beta r}\right) - \eta\frac{(1-\gamma)}{\beta r}$, then it is further restricted by the parameters of the HARA utility function, and features the same arbitrary constant used for the multiplicative transformation.

where $a, b, g,$and h are, at most, functions of time, if and only if symmetry 4 (the scaling symmetry) is satisfied.

Note that in the above corollary no reference is made to any specific form of the utility function. The proof of this corollary does not necessitate a specific solution to the HJB equation in the model and in any case it is impossible to give a precise solution when the utility function is not specified.

This means that wealth scale invariance implies linear optimal solutions to the control variables (C^*, w^*) and linear optimal rules imply scale invariance. Scale invariance determines the relevant linear parameters of optimal behavior.

(iii) If we know that we can compare the outcome of two different consumers as a linear function of the ratio of their wealth stocks, then *necessarily* the utility function of the agents is of the HARA form. This can also be stated as follows: if the outcome of two different consumers cannot be compared as a linear function of the ratio of their wealth, then their utility function is not HARA. We can state this even without solving the model explicitly, as it emerges from the analysis of the symmetries, i.e., the invariance properties of the HJB equation. This result stems from the formulation of symmetry 4 whereby if J is a solution then a linear function of J is a solution and from the 'if and only if' property discussed above.

4.3 Linear Connections: Implications for Risk Aversion

A utility function $U(C)$ is said to be HARA *iff* it has a risk tolerance $T(C)$ which is a linear function of its argument. That is:

$$T(C) = \frac{1}{A(C)} = \frac{C}{1-\gamma} + \frac{\eta}{\beta} \tag{32}$$

where absolute risk aversion $A(C)$ is given by:

$$A(C) = -\frac{U''(C)}{U'(C)}$$

Thus, there are three linear formulations here which are inter-connected:

a. Scale invariance is established through linear transformations of wealth; see the transformation $e^{-k}\{W + \frac{(1-\gamma)\eta}{\beta r}\} - \eta\frac{(1-\gamma)}{\beta r}$ in Symmetry 4 above and the ensuing discussion.

b. Optimal consumption C^* and portfolio shares w^* are linear functions of wealth; see Corollary 3 above.

c. Risk aversion is such that its reciprocal, risk tolerance, is linear in consumption, as in equation (32).

Symmetry 4 establishes the equivalence between these three linear formulations, all with an iff property: scaling invariance – as in point a – generates HARA, which yields optimal behavior, as in point b, and which risk aversion is defined in point c.

Note that the functional equation (30), which holds true *iff* there is scale invariance, is the same as the functional equation which has to be satisfied by the utility function in order for it to feature linear risk tolerance (equation (32)). This fundamental functional equation is implicit in the model itself and the way to uncover it is through the calculation of the Lie symmetries.

One can then interpret the results above also as follows: a requirement of risk tolerance to be linear in consumption as an "economic fundamental" means that utility has to be HARA, and through the equivalence implied by Symmetry 4, that the indirect utility function be scale invariant.

5 Conclusions

The analysis has derived the set of all the transformations that leave the optimal solution invariant in the consumer-investor problem analyzed by Merton (1969, 1971). In particular, we have derived HARA utility as an inherent feature, a restriction that applies when invariance of the optimal solution is to obtain under multiplicative transformations of wealth. Doing so, we showed that wealth scale invariance implies linear optimal solutions and linear risk tolerance.

While undertaking this analysis, the paper has demonstrated the use of Lie symmetries as a powerful tool to deal with economic optimization problems. Such invariance restrictions are fundamental to economic analysis. The analysis demonstrates how economic conclusions flow from economic assumptions, as opposed to arbitrary functional form assumptions. Lie symmetries techniques are a key tool for finding the set of functional forms implied by a given economic assumption.

There are likely to be many other optimization problems that would yield restrictions of the type explored here. Consider two examples: first, the problem, examined by Lucas and Stokey (1984) and subsequent literature, of modelling optimal paths in an economy with heterogenous agents and growth, is likely to be amenable to such analysis. A much more recent treatment of this kind of problem has been undertaken in the context of DSGE models and the relationship between representative agent models and heterogeneous agent economies (see Chang and Kim (2007) and An, Chang and Kim (2009) for examples). Symmetries can provide conditions for aggregator functions and restrictions on the utility functions. Second, for problems with non-constant rates of time preference, such as the one analyzed by Barro (1999), symmetries can provide restrictions on the time preference function. Hence, rather than assume certain properties of such functions, these could be derived using the tools presented above.

Lie symmetries may also be used to characterize solutions when these cannot be represented in closed-form. However, as far as we can see, there is no general principle here which could be formulated. These issues need to be tackled on a case by case basis. In some the symmetries will yield trivial structural characterizations, while in others they could generate significant insights.

References

[1] An, Sungbae, Yongsung Chang, and Sun-Bin Kim, 2009. “Can a Representative Agent Model Represent a Heterogeneous Agent Economy?” **American Economic Journal – Macroeconomics** 1, 29-54.

[2] Barro, Robert J., 1999. “Ramsey Meets Laibson in the Neoclassical Growth Model,” **The Quarterly Journal of Economics** 114, 4, 1125-1152.

[3] Boyd III, John H., 1990. “Symmetries, Dynamic Equilibria and the Value Function” in R.Sato and R.V. Ramachandran (eds.), **Conservation Laws and Symmetries**, Kluwer Academic Publishers, Boston, 225-260.

[4] Brunnermeier, Markus K. and Stefan Nagel, 2008. “Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals’ Asset Allocation,”**American Economic Review** 98:3, 713-736.

[5] Chang, Yongsung, and Sun-Bin Kim, 2007. “Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations,” **American Economic Review** 97, 1939-1956.

[6] Duffie, Darrell, 2003. “Intertemporal Asset Pricing Theory” in G.M. Constantinides, M. Harris and R. Stulz (eds.) **Handbook of the Economics of Finance** Vol. 1B, 639–742.

[7] Lucas, Robert E., 1976. “Econometric Policy Evaluation: A Critique,” **Carnegie-Rochester Conference on Public Policy**, 1, 19-46.

[8] Lucas, Robert E., and Nancy L. Stokey, 1984. “Optimal Growth with Many Consumers,” **Journal of Economic Theory** 32, 139-171.

[9] Merton, Robert C., 1969, "Lifetime Portfolio Selection Under Uncertainty: The Continuous Time Case," **Review of Economics and Statistics** 51, 247-57.

[10] Merton, Robert C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous Time Model," **Journal of Economic Theory** 3, 373-413.

[11] Merton, Robert C., 1990. **Continuous Time Finance**, Blackwell, Cambridge, MA.

[12] Olver, Peter J., 1993. **Applications of Lie Groups to Differential Equations**, 2nd edition, Springer-Verlag, New York.

[13] Sato, Ryuzo, 1981. **Theory of Technical Change and Economic Invariance: Application of Lie Groups**, Academic Press, New York. Reprinted with amendments, 1999, Edward Elgar Publishing Ltd., Cheltenham.

[14] Sato, Ryuzo, and Rama V.Ramachandran (eds.), 1990. **Conservation Laws and Symmetries**, Kluwer Academic Publishers, Boston.

[15] Sato, Ryuzo, and Rama V. Ramachandran, 1997. **Symmetry and Economic Invariance: An Introduction**, Kluwer Academic Publishers, Boston.

[16] Weil, Philippe, 1989. "The Equity Premium Puzzle and the Risk Free Rate Puzzle," **Journal of Monetary Economics** 24, 401-421.

[17] Weil, Philippe, 1990. "Nonexpected Utility in Macroeconomics," **The Quarterly Journal of Economics,** 105,1, 29-42.