Concave Consumption Function and Precautionary Wealth Accumulation*

Richard M. H. Suen†

Abstract

This paper examines the theoretical foundations of precautionary saving behavior in a canonical life-cycle model where consumers face uninsurable idiosyncratic labor income risk and borrowing constraints. We begin by characterizing the consumption function of individual consumers. We show that the consumption function is concave when the utility function has strictly positive third derivative and the inverse of absolute prudence is a concave function. These conditions encompass all HARA utility functions with strictly positive third derivative as special cases. We then show that when consumption function is concave, a mean-preserving increase in income risks would encourage wealth accumulation at both the individual and aggregate levels.

Keywords: Consumption function, borrowing constraints, precautionary saving

JEL classification: D81, D91, E21.

*The author thanks Alan Binder, Christopher Carroll, Jang-Ting Guo, Miles Kimball, Philipp Kircher (the editor), Carine Nourry, three anonymous referees, the seminar participants at the University of Connecticut and UC Riverside, and the conference participants at the 2011 Symposium of SNDE, the 2011 T2M Conference, the 2011 North American Summer Meetings of the Econometric Society and the 2014 Midwest Economic Theory Conference for helpful comments and suggestions.

†Department of Economics, 365 Fairfield Way, Unit 1063, University of Connecticut, Storrs CT 06269-1063, United States. Email: *richard.suen@uconn.edu.* Phone: 1 860 486 4368. Fax: 1 860 486 4463.

1 Introduction

This paper examines the theoretical foundations of precautionary saving behavior in a canonical life-cycle model where consumers face uninsurable idiosyncratic labor income risk and borrowing constraints. Within this framework, we seek general conditions on the utility function under which a mean-preserving increase in labor income risk will increase aggregate savings. This study is motivated by the following observations.

Precautionary saving behavior has long been the subject of both empirical and theoretical research. There is now a large number of empirical studies that find a positive relationship between wealth accumulation (or consumption growth) and uncertainty using household-level data.[1] On the theory side, the standard incomplete markets model is now a common framework for analyzing macroeconomic issues.[2] This type of model is built upon the premise that individual's precautionary motive of saving is an important driving force behind aggregate wealth accumulation and economic inequality. Thus, understanding the theoretical underpinning of precautionary wealth accumulation is of fundamental importance to this line of research. Since the work of Leland (1968), Sibley (1975), Miller (1976) and Mendelson and Amihud (1982), it has been known that if individuals have time-separable preferences and if the third derivative of their (per-period) utility function is positive, then individual savings will increase in response to greater income risk. However, as pointed out by Huggett (2004), this does not necessarily imply that aggregate savings will also increase. This is due to the following reason. In models with idiosyncratic uncertainty and imperfect insurance, aggregate savings are formed by integrating individual savings over the cross-sectional distribution of individual states. An increase in income risk will not only affect the level of individual savings, it will also make the distribution of individual states more dispersed, and this can either raise or lower the level of aggregate savings. Huggett (2004) shows that if, in addition to a positive third derivative, individuals' consumption function (i.e., the policy function for consumption) is increasing and concave then greater income risk will increase aggregate savings.[3] This result has at least two important implications. First, it suggests that caution needs to be exercised when one attempts to infer the existence or importance of aggregate

[1] See, for instance, Carroll and Samwick (1997, 1998), Lusardi (1998), Gourinches and Parker (2002), Parker and Preston (2005), and the papers surveyed in Browning and Lusardi (1996, Section 5.3).

[2] An extensive survey on this literature is provided in Heathcote *et al.* (2009).

[3] More precisely, Huggett shows that if the level of individual savings increases with risk, and if the policy function for savings is increasing and convex, then greater income risk will increase aggregate savings. In both Huggett's and our model, a convex policy function for savings is equivalent to a concave consumption function.

precautionary savings from household-level evidence. Second, and more important to the focus of our paper, this result suggests that one way to better understand aggregate precautionary savings is to identify the conditions under which consumption function is concave.[4]

The concavity of consumption function has been previously examined by other authors. Earlier studies, such as Zeldes (1989), Deaton (1991) and Carroll (1992, 1997), solve the "buffer-stock" saving model using numerical methods and find that the consumption function is concave under constant-relative-risk-aversion utility. Carroll and Kimball (1996) provide the first formal proof that consumption function is concave under any HARA utility function with strictly positive third derivative.[5] Huggett (2004) and Carroll and Kimball (2005) extend this result to different model environments, but maintain the HARA assumption. To the best of our knowledge, there are no existing studies that have analyzed this issue without invoking this assumption. Thus, our current knowledge of concave consumption function and aggregate precautionary savings is still limited to one particular class of utility functions. This raises the question of whether the rather restrictive assumption of HARA utility is necessary for these results.[6] The present study is intended to answer this question.

In this paper, we consider an economy in which a large number of *ex ante* identical consumers face uninsurable idiosyncratic labor income risk and borrowing constraints over a finite lifetime. The exogenous income process consists of a permanent shock and a purely transistory shock. In the first part of the paper, we focus on the intertemporal consumption-saving problem faced by a typical consumer and provide a detailed characterization of the policy functions. Unlike existing studies, we do not confine our attention to HARA utility functions. Instead, we seek to provide a set of general conditions on the utility function under which the consumption function is concave. In the second part of the paper, we explore the connection between concave consumption function and precautionary saving. In particular, we examine how changes in the riskiness of the permanent shock and the purely transitory shock will affect wealth accumulation at both the individual and aggregate levels. Similar

[4] The concavity of consumption function is also of interest in its own right. Specifically, this property implies that more affluent consumers have a lower propensity to consume than less affluent ones. This prediction is consistent with the empirical evidence reported in Johnson *et al.* (2006) and Jappelli and Pistaferri (2014).

[5] HARA is the acronym for hyperbolic absolute risk aversion. Examples of HARA utility functions include the constant-absolute-risk-aversion (CARA) utility functions, the constant-relative-risk-aversion (CRRA) utility functions and the quadratic utility functions.

[6] The HARA assumption has a number of strong implications on consumers' preferences. For instance, this assumption is equivalent to assuming that consumers have linear absolute risk tolerance. There is no empirical evidence concerning the linearity or the curvature of absolute risk tolerance. In the theory of consumer preferences, the usual axioms do not imply a linear absolute risk tolerance. On the contrary, a growing number of studies show that the *nonlinearity* of absolute risk tolerance has important implications in both deterministic and stochastic models. See, for instance, Gollier (2001a), Gollier and Zeckhauser (2002), Ghiglino (2005) and Ghiglino and Venditti (2011).

to Huggett (2004), we focus on the effects of greater income risk on the cross-sectional average level of asset holdings at different stages of the life cycle (i.e., the average life-cycle profile of wealth).

This paper has two major findings. The first one is a novel set of conditions under which consumption function is concave. Specifically we show that if, in addition to a strictly positive third derivative of the utility function, the inverse of absolute prudence is a concave function, then consumption function is concave at every stage of the life cycle.[7] We also show that these conditions include as special case all HARA utility functions with strictly positive third derivative, thus proving that the HARA assumption is *not necessary* for the concavity result. When comparing to Huggett (2004), our first result is more general in two aspects. First, Huggett only considers purely transitory shock, while the present study takes into account both permanent and transitory shocks. This dichotomy between permanent and transitory shocks is now commonly used in quantitative studies.[8] We show that distinguishing between these two types of shocks also brings some new insights to the theoretical results. Second, Huggett only focuses on HARA utility functions. We show that the concavity result can be obtained under more general conditions. Our second major finding states that, when the consumption function at every stage of the life cycle is concave, a mean-preserving increase in either the permanent shock or the purely transitory shock will promote wealth accumulation at both the individual and aggregate levels. This result not only establishes the connection between concave consumption function and precautionary wealth accumulation, it also has other implications regarding saving behavior in the presence of multiple risks. One implication is that, when consumption functions are concave, the introduction of an independent "background" risk in some distant future (say, in period $t+n$, where $n>1$) will amplify the precautionary saving motive in period t.[9] Similar results have been previously reported by Blundell and Stoker (1999). Specifically, these authors consider a three-period model in which consumers with CRRA utility face income uncertainty in the last two periods. Through a series of numerical examples, they show that an increase in income risk in the third period (the distant future) will reinforce the precautionary saving motive in the first period (see, for instance,

[7] The term "absolute prudence" is first introduced by Kimball (1990). For a thrice differentiable utility function $u(\cdot)$, the coefficient of absolute prudence is defined as $\Pi(c) \equiv -u'''(c)/u''(c)$.

[8] See, for instance, Zeldes (1989), Carroll (1992, 1997, 2011), Ludvigson (1999), Gourinchas and Parker (2002) and Storesletten *et al.* (2004a), among many others. Empirical studies on household earnings dynamics, such as Abowd and Card (1989), Storesletten *et al.* (2004b), and Moffitt and Gottschalk (2011), show that this specification fits the data well.

[9] There is now a large number of studies which explore the effects of background risks on consumer behavior. Most of these studies, however, adopt an atemporal environment and and thus do not consider the effects of background risks on precautionary saving behavior. For a textbook treatment of this literature, see Gollier (2001b, Chapters 8 and 9).

their scenarios 1.1 and 3.1). Our results in Section 4 provide a formal theoretical foundation for their findings and show that similar results can be obtained in a more general environment.

This rest of this paper is organized as follows. Section 2 describes the model environment and establishes some intermediate results. Section 3 establishes the concavity of the consumption function and contrasts our result to those in the existing studies. Section 4 explores the connection between concave consumption function and precautionary saving behavior. Section 5 provides some concluding remarks.

2 The Model

Consider an economy inhabited by a continuum of *ex ante* identical consumers. The size of population is normalized to one. Each consumer faces a planning horizon of $(T+1)$ periods, where T is finite. The consumers have preferences over random consumption paths $\{c_t\}_{t=0}^{T}$ which can be represented by

$$E_0\left[\sum_{t=0}^{T}\beta^t u\left(c_t\right)\right], \tag{1}$$

where $\beta \in (0,1)$ is the subjective discount factor and $u(\cdot)$ is the (per-period) utility function. The domain of the utility function is given by $\mathcal{D} = [\underline{c}, \infty)$, where $\underline{c} \geq 0$ can be interpreted as a minimum consumption requirement.[10] The function $u : \mathcal{D} \rightarrow [-\infty, \infty)$ is once continuously differentiable, strictly increasing and strictly concave.[11] There is no restriction on $u(\underline{c})$ which means the utility function can be either bounded or unbounded below.

In each period $t \in \{0, 1, ..., T\}$, each consumer receives a random amount of labor endowment e_t, which they supply inelastically to work. Labor income in period t is given by we_t, where $w > 0$ is a constant wage rate. The stochastic labor endowment is determined by $e_t = \widetilde{e}_t \varepsilon_t$, where $\widetilde{e}_t$ is a permanent shock and ε_t is a purely transitory shock. The initial value of the permanent shock $\widetilde{e}_0 > 0$ is exogenously given and is identical across consumers. The random variable $\widetilde{e}_t$ then evolves according to $\widetilde{e}_t = \widetilde{e}_{t-1}\nu_t$, where ν_t is drawn from a compact interval $\Lambda \equiv [\underline{\nu}, \overline{\nu}]$, with $\overline{\nu} > \underline{\nu} > 0$, according to the distribution $L_t(\cdot)$. Similarly, the transitory shock ε_t is drawn from a compact interval $\Xi \equiv [\underline{\varepsilon}, \overline{\varepsilon}]$, with

[10] This specification allows for utility functions that are not defined at $c = 0$. One example is the Stone-Geary utility function which belongs to the HARA class and features a minimum consumption requirement. All the results in this paper remain valid if we set $\underline{c} = 0$.

[11] None of the results in this section require higher order differentiability of the utility function. This property is only needed in later sections.

$\overline{\varepsilon} > \underline{\varepsilon} > 0$, according to the distribution $G_t(\cdot)$.[12] Both $L_t(\cdot)$ and $G_t(\cdot)$ are well-defined distribution functions, which means they are both nondecreasing, right-continuous and satisfy the conditions

$$\lim_{\nu \to \underline{\nu}} L_t(\nu) = \lim_{\varepsilon \to \underline{\varepsilon}} G_t(\varepsilon) = 0 \quad \text{and} \quad \lim_{\nu \to \overline{\nu}} L_t(\nu) = \lim_{\varepsilon \to \overline{\varepsilon}} G_t(\varepsilon) = 1.$$

The two random variables ε_t and ν_t are independent of each other, across time and across consumers.

Under this specification, knowledge on both $\widetilde{e}_t$ and ε_t is required to determine the distribution of e_{t+1}. Thus, an individual's state in period t includes $z_t \equiv (\widetilde{e}_t, \varepsilon_t)$. The above assumptions imply that $z_t \equiv (\widetilde{e}_t, \varepsilon_t)$ follows a Markov process with compact state space $Z_t \equiv \Delta_t \times \Xi$, where $\Delta_t \equiv \left[\underline{\xi}_t, \overline{\xi}_t\right]$ contains all possible realizations of $\widetilde{e}_t$.[13] Let $(Z_t, \mathcal{Z}_t)$ be a measurable space and $Q_t : (Z_t, \mathcal{Z}_t) \to [0,1]$ be the transition function of the Markov process at time t. For any $z = (\widetilde{e}, \varepsilon) \in Z_t$ and for any $B \subseteq Z_{t+1}$,

$$Q_t(z, B) \equiv \Pr\{(\nu_{t+1}, \varepsilon_{t+1}) \in \Lambda \times \Xi : (\widetilde{e}\nu_{t+1}, \varepsilon_{t+1}) \in B\}.$$

It is straightforward to show that the transition function satisfies the Feller property in every period.

In light of the uncertainty in labor income, consumers can only self-insure by borrowing or lending a single risk-free asset. The gross return from this asset is $(1+r) > 0$.[14] Let a_t denote asset holdings in period t. A consumer is said to be in debt if a_t is negative. In each period, the consumers are subject to the budget constraint:

$$c_t + a_{t+1} = we_t + (1+r)\,a_t, \tag{2}$$

and the borrowing constraint: $a_{t+1} \geq -\underline{a}_{t+1}$. The parameter $\underline{a}_{t+1} \geq 0$ represents the maximum amount of debt that a consumer can owe in period $t+1$. In life-cycle models, the borrowing limit in the terminal period is typically set to zero, which means the consumers cannot die in debt. But the borrowing limit in all other periods can be different from zero. Throughout this paper, we maintain

[12] The age-dependent nature of $L_t(\cdot)$ and $G_t(\cdot)$ can be used to capture any age-specific differences in earnings across consumers. Hence, there is no need to include a separate life-cycle component in e_t.

[13] For any $t \in \{0, 1, ..., T\}$, the upper and lower bounds of Δ_t are given by $\overline{\xi}_t \equiv \widetilde{e}_0 \overline{\nu}^t$ and $\underline{\xi}_t \equiv \widetilde{e}_0 \underline{\nu}^t$, respectively.

[14] Existing studies typically assume that w and r are known constants. See, for instance, Deaton (1991), Carroll (1992, 1997) and Huggett (2004). This assumption is not essential for our results. Specifically, all of our proofs can be easily modified to handle any price sequences $\{w_t, r_t\}_{t=0}^{T}$ that are deterministic, strictly positive and finite-valued.

the following assumptions on the borrowing limits: $\underline{a}_t \geq 0$, for all t, $\underline{a}_{T+1} = 0$ and

$$w\underline{e}_t - (1+r)\,\underline{a}_t + \underline{a}_{t+1} > \underline{c}, \qquad \text{for all } t \geq 0, \tag{3}$$

where $\underline{e}_t \equiv \underline{\xi}_t \times \underline{\varepsilon}$ is the lowest possible value of e_t. The intuitions of (3) are as follows. Consider a consumer who faces the worst possible state in period t, i.e., $a_t = -\underline{a}_t$ and $e_t = \underline{e}_t$. The highest attainable consumption in this particular state is $c_t = w\underline{e}_t - (1+r)\,\underline{a}_t + \underline{a}_{t+1}$. Condition (3) thus ensures that the minimum consumption requirement can be met even in the worst possible state. The same condition also ensures that any debt in period t can be fully repaid even if a consumer receives the lowest labor income in all future time periods, i.e.,

$$\sum_{j=0}^{T-t} \frac{w\underline{e}_{t+j} - \underline{c}}{(1+r)^j} - (1+r)\,\underline{a}_t > 0, \qquad \text{for all } t \geq 0.$$

2.1 Consumers' Problem

Given the prices w and r, the consumers' problem is to choose sequences of consumption and asset holdings, $\{c_t, a_{t+1}\}_{t=0}^{T}$, so as to maximize the expected lifetime utility in (1), subject to the budget constraint in (2), the minimum consumption requirement $c_t \geq \underline{c}$, the borrowing constraint $a_{t+1} \geq -\underline{a}_{t+1}$ for all t, and the initial conditions: $a_0 \geq -\underline{a}_0$ and $\widetilde{e}_0 > 0$. In the present study, we focus on consumers whose rate of time preference $\rho \equiv 1/\beta - 1$ is no less than the market interest rate, so that $\beta\,(1+r) \leq 1$.[15]

Define a sequence of assets $\{\overline{a}_t\}_{t=0}^{T}$ according to

$$\overline{a}_{t+1} = w\overline{e}_t + (1+r)\,\overline{a}_t - \underline{c}, \qquad \text{for all } t \geq 0,$$

where $\overline{e}_t \equiv \overline{\xi}_t \times \overline{\varepsilon}$ is the highest possible value of labor endowment and $\overline{a}_0 = a_0$. This sequence specifies the amount of assets that can be accumulated if a consumer receives the highest possible labor income and consumes only the minimum requirement $\underline{c}$ in every period. Since $(1+r) > 0$, this sequence is

[15] This assumption is standard in "buffer-stock" saving models. See, for instance, Deaton (1991) and Carroll (1992, 1997).

monotonically increasing and bounded above by

$$\overline{a}_T \equiv (1+r)^T a_0 + \sum_{j=0}^{T-1} (1+r)^{T-1-j} (w\overline{e}_j - \underline{c}),$$

which is finite as T is finite. Any feasible sequence of assets $\{a_t\}_{t=0}^T$ must be bounded above by $\{\overline{a}_t\}_{t=0}^T$. Hence, the state space of asset in every period t can be restricted to the compact interval $A_t = [-\underline{a}_t, \overline{a}_t]$.

In any given period, a consumer's state is summarized by $s = (a, z)$, where a denotes his current asset holdings, and $z \equiv (\widetilde{e}, \varepsilon)$ is the current realization of the shocks.[16] The set of all possible states in period t is given by $S_t = A_t \times Z_t$ which is compact.

Define a set of value functions $\{V_t\}_{t=0}^T$, $V_t : S_t \to [-\infty, \infty]$ for each t, recursively according to

$$V_t(a,z) = \max_{c \in [\underline{c}, x(a,z)+\underline{a}_{t+1}]} \left\{ u(c) + \beta \int_{Z_{t+1}} V_{t+1}[x(a,z) - c, z'] \, Q_t(z, dz') \right\}, \tag{P1}$$

where $x(a,z) \equiv we(z) + (1+r)a$ and $e(z)$ is the level of labor endowment under $z \equiv (\widetilde{e}, \varepsilon)$. The variable $x(a,z)$ is often referred to as *cash-in-hand* in the existing studies. In the terminal period, the value function is given by

$$V_T(a,z) = u[we(z) + (1+r)a], \qquad \text{for all } (a,z) \in S_T.$$

Define a set of optimal policy correspondences for consumption $\{g_t\}_{t=0}^T$ according to

$$g_t(a,z) \equiv \arg\max_{c \in [\underline{c}, x(a,z)+\underline{a}_{t+1}]} \left\{ u(c) + \beta \int_{Z_{t+1}} V_{t+1}[x(a,z) - c, z'] \, Q_t(z, dz') \right\}, \tag{4}$$

for all $(a,z) \in S_t$ and for all t. Given $g_t(a,z)$, the optimal choices of a_{t+1} are given by

$$h_t(a,z) \equiv \left\{ a' : a' = x(a,z) - c, \text{ for some } c \in g_t(a,z) \right\}. \tag{5}$$

The main properties of the value functions and the optimal policy correspondences are summarized

[16] For the results in this section, the distinction between $\widetilde{e}$ and ε is immaterial. Hence, we express individual state as $s = (a, z)$ instead of $s = (a, \widetilde{e}, \varepsilon)$.

in Lemma 1.[17] The first part of the lemma states that the value function in every period t is bounded and continuous on S_t. This is true even if $u(\cdot)$ is unbounded below. Boundedness of the value function ensures that the conditional expectation in (P1) is well-defined. Continuity of the objective function in (P1) ensures that $g_t(\cdot)$ and $h_t(\cdot)$ are non-empty and upper hemicontinuous. The second part of the lemma establishes the strict monotonicity and strict concavity of $V_t(\cdot, z)$. The latter property implies that $g_t(\cdot, z)$ and $h_t(\cdot, z)$ are single-valued functions for all $z \in Z_t$. Part (iii) establishes the differentiability of $V_t(\cdot, z)$. Specifically, it shows that $V_t(\cdot, z)$ is not only differentiable in the interior of A_t, it is also (right-hand) differentiable at the lower endpoint $-\underline{a}_t$. This property is important because, as is well-known in this literature, a consumer may choose to exhaust the borrowing limit in some states.[18] In other words, the consumer's problem may have corner solutions in which $h_{t-1}(a, z) = -\underline{a}_t$ for some $(a, z) \in S_{t-1}$. Thus, the first-order condition of (P1) has to be valid even when $h_{t-1}(a, z) = -\underline{a}_t$ and this requires $V_t(\cdot, z)$ to be right-hand differentiable at $a = -\underline{a}_t$.[19] Part (iv) of the lemma establishes the Euler equation for consumption. Part (v) states that the consumption function in every period is a strictly increasing function.

Lemma 1 *The following results hold for all* $t \in \{0, 1, ..., T\}$.

(i) The value function $V_t(a, z)$ *is bounded and continuous on* S_t.

(ii) For every $z \in Z_t$, $V_t(\cdot, z)$ *is strictly increasing and strictly concave on* A_t.

(iii) For every $z \in Z_t$, $V_t(\cdot, z)$ *is continuously differentiable on* $[-\underline{a}_t, \overline{a}_t)$ *and the derivative of this function is given by*

$$\frac{\partial V_t(a, z)}{\partial a} = (1 + r)\, u'[g_t(a, z)], \qquad \text{for all } a \in [-\underline{a}_t, \overline{a}_t).$$

[17] Most of these properties are standard and have been proved elsewhere, hence the proof of Lemma 1 is omitted. The complete set of proofs can be found in the working paper version of this paper.

[18] A formal proof of this result can be found in Mendelson and Amihud (1982), Aiyagari (1994), Huggett and Ospina (2001) and Rabault (2002).

[19] Note that the standard result in Stokey, Lucas and Prescott (1989) Theorem 9.10 only establishes the differentiability of the value function in the interior of the state space. Thus, additional effort is needed to establish this result. Another related point is that $V_t(\cdot, z)$ is differentiable on $[-\underline{a}_t, \overline{a}_t)$ even if the policy function $g_t(\cdot, z)$ is not differentiable at the point where the borrowing constraint becomes binding. For a more detailed discussion on this, see Schechtman (1976) p.221-222.

(iv) For all $(a, z) \in S_t$, the policy functions $g_t(a, z)$ and $h_t(a, z)$ satisfy the Euler equation

$$u'[g_t(a, z)] \geq \beta (1 + r) \int_{Z_{t+1}} u'\left[g_{t+1}\left(h_t(a, z), z'\right)\right] Q_t\left(z, dz'\right), \tag{6}$$

with equality holds if $h_t(a, z) > -\underline{a}_{t+1}$.

(v) The consumption function $g_t(\cdot)$ is a strictly increasing function.

3 Concavity of Consumption Function

In this section, we provide a set of conditions on $u(\cdot)$ such that the consumption function at every stage of the life cycle exhibits concavity. From this point onwards, we will focus on utility functions that are four-times differentiable, strictly increasing and strictly concave. The main result of this section is summarized in Theorem 1 which covers two groups of utility functions: (i) quadratic utility functions, or those with $u'''(c) = 0$ throughout its domain, and (ii) utility functions with strictly positive third derivative and the *inverse of absolute prudence* $I(c) \equiv -u''(c)/u'''(c)$ is a concave function. As we will see later, the second group of utility functions include all HARA utility functions with strictly positive third derivative as special case.

Theorem 1 *Suppose the utility function $u(\cdot)$ is four-times differentiable, strictly increasing, strictly concave and satisfies one of the following conditions:* (i) $u'''(c) = 0$, *or* (ii) $u'''(c) > 0$ *and the inverse of absolute prudence $I(c) \equiv -u''(c)/u'''(c)$ is concave throughout the domain of $u(\cdot)$. Then the following results hold for all $t \in \{0, 1, ..., T\}$.*

(i) *For any $\varepsilon \in \Xi$, the consumption function $g_t(a, \widetilde{e}, \varepsilon)$ is concave in $(a, \widetilde{e})$.*

(ii) *For any $\widetilde{e} \in \Delta_t$, the consumption function $g_t(a, \widetilde{e}, \varepsilon)$ is concave in (a, ε).*

The proof of Theorem 1 can be found in the Appendix. Here we provide a heuristic discussion on the main ideas of the proof. As shown in Lemma 1, the consumption functions must satisfy the Euler equation in (6). This equation essentially defines an operator (often referred to as the "Euler operator") which maps the consumption function in period $t + 1$ to that in period t. Thus, starting from $g_T(\cdot)$, one can derive the consumption function in all preceding periods recursively using this operator. When the utility function is quadratic, the Euler operator is a linear mapping. This, together with the fact that $g_T(a, \widetilde{e}, \varepsilon)$ is linear in $(a, \widetilde{e})$ and in (a, ε), implies that the consumption function in all preceding periods are (piecewise) linear in $(a, \widetilde{e})$ and in (a, ε).[20] The main task of the proof is to show that, for the second group of utility functions, concavity is preserved by the Euler operator, i.e., if $g_{t+1}(\cdot)$ is concave in $(a, \widetilde{e})$ and in (a, ε), then $g_t(\cdot)$ will also have these properties.

Three remarks about Theorem 1 are in order. First, the theorem states that the consumption functions are jointly concave in current asset holdings and one of the shocks, when the other is held constant. In particular, $g_t(a, \widetilde{e}, \varepsilon)$ is *not* jointly concave in $(a, \widetilde{e}, \varepsilon)$. This happens because the stochastic labor endowment $e \equiv \widetilde{e}\varepsilon$ is not jointly concave in $(\widetilde{e}, \varepsilon)$. Consequently, the consumption function in the terminal period, which serves as the starting point of the backward induction process, is not jointly concave in $(a, \widetilde{e}, \varepsilon)$.[21] Second, concavity in $(a, \widetilde{e})$ implies concavity in (a, ε), but the converse is not true in general. Thus, it is essential to distinguish between permanent shock and purely transitory shock. Third, Theorem 1 has the following implications regarding the propensity of consumption: Consider two age-t consumers with the same (a, ε) but different values of $\widetilde{e}$, say $\widetilde{e}_1 > \widetilde{e}_2$. Then the one with a higher labor income $(w\widetilde{e}_1\varepsilon)$ will have a lower propensity to consume than the one with a lower labor income $(w\widetilde{e}_2\varepsilon)$. Formally, this means for any $\varkappa > 0$, we have

$$\frac{g_t(a, \widetilde{e}_1 + \varkappa, \varepsilon) - g_t(a, \widetilde{e}_1, \varepsilon)}{w\varepsilon\varkappa} \leq \frac{g_t(a, \widetilde{e}_2 + \varkappa, \varepsilon) - g_t(a, \widetilde{e}_2, \varepsilon)}{w\varepsilon\varkappa}.$$

Similarly, holding age and $(a, \widetilde{e})$ constant, a consumer with a higher value of ε will have a lower propensity to consume.

We now compare our results to those in the existing literature. As mentioned above, existing studies typically confine their attention to the HARA class of utility functions. A utility function is called HARA if there exists $(\alpha, \gamma) \in \mathbb{R}^2$ such that

$$-\frac{u''(c)}{u'(c)} = \frac{1}{\alpha + \gamma c} \geq 0, \qquad \text{for all } c \in \mathcal{D}. \tag{7}$$

[20] If the borrowing constraint is never binding, then the consumption function in any period $t < T$ is linear in $(a, \widetilde{e})$ and in (a, ε) when the utility function is quadratic. If the borrowing constraint is binding in some states, then the consumption function is kinked and piecewise linear in these variables.

[21] The assumption of $e \equiv \widetilde{e}\varepsilon$ also implies that the graph of the budget correspondence

$$\mathcal{B}_t(a, \widetilde{e}, \varepsilon) \equiv \left\{c : \underline{c} \leq c \leq w\widetilde{e}\varepsilon + (1 + r)a + \underline{a}_{t+1}\right\},$$

which contains the graph of $g_t(a, \widetilde{e}, \varepsilon)$, is not a convex set itself. On the other hand, if the stochastic labor endowment e is determined by $e = f(\widetilde{e}, \varepsilon)$, where f is jointly concave in its arguments, then the graph of $\mathcal{B}_t(a, \widetilde{e}, \varepsilon)$ is a convex set and the consumption functions are jointly concave in $(a, \widetilde{e}, \varepsilon)$.

HARA utility functions are smooth functions, which means they are *infinitely* differentiable. Examples of HARA utility functions include the quadratic utility functions, the exponential utility functions, the CRRA utility functions and the Stone-Geary utility functions. All these utility functions, except the quadratic ones, have strictly positive third derivative. Differentiating both sides of (7) with respect to c gives

$$\frac{u'(c)\,u'''(c)}{\left[u''(c)\right]^2} = 1+\gamma, \qquad \text{for all } c \in \mathcal{D}. \tag{8}$$

Thus, the third derivative is strictly positive if and only if $\gamma > -1$. Combining (7) and (8) gives

$$I(c) \equiv -\frac{u''(c)}{u'''(c)} = \frac{\alpha+\gamma c}{1+\gamma}.$$

This shows that the inverse of absolute prudence for any HARA utility function with strictly positive third derivative is a linear function.[22] The conditions stated in Theorem 1 thus contain these utility functions as special case.[23] When comparing to Huggett (2004) Lemma 1, our results are more general in two ways. First, Huggett only considers serially independent labor income shocks, whereas we consider both permanent and purely transistory shocks. Second, Huggett proves that consumption functions are concave for exponential and CRRA utility functions. We establish the same result under a general specification of the utility function, without confining ourselves to the HARA class. This approach proves useful in developing new insights on this topic. Specifically, our Theorem 1 shows that the HARA assumption is not required for the concavity result. Instead, it is determined by a more fundamental property of consumer preferences, which is the shape of the inverse of absolute prudence.[24] Our results also complement those in Huggett and Vidon (2002). Using specific numerical examples, these authors show that a strictly positive $u'''(\cdot)$ alone is not enough to generate concave consumption functions. They, however, have not provided any additional conditions that are needed for this result. Our Theorem 1 suggests that the concavity of $I(\cdot)$ may be the missing piece of the puzzle.

[22] In general, the HARA assumption implies that all nth order index of absolute risk aversion, i.e., $A_n(c) \equiv -u^{(n+1)}(c)/u^{(n)}(c)$, $n \in \{1,2,...\}$, is hyperbolic. Thus, the inverse of each $A_n(c)$ is a linear function in c.

[23] For HARA utility functions with strictly positive third derivative, we can sharpen the results in Theorem 1 and show that the consumption function in any period $t < T$ is *strictly* concave in $(a, \widetilde{e})$ and in (a, ε) when the borrowing constraint is not binding. The details of these can be found in the working paper version of this paper.

[24] Two other studies have explored the implications of a concave $I(\cdot)$ in other contexts. Gollier (2001a) shows that if the inverse of absolute prudence is a concave function, then an increase in wealth inequality in a complete market economy will reduce the equilibrium risk-free rate. Mazzocco (2004) shows that efficient risk sharing within a household will lower savings if and only if the inverse of absolute prudence is concave.

4 Precautionary Wealth Accumulation

We now explore the implications of concave consumption function on precautionary wealth accumulation, both at the individual and aggregate levels. Recall the economy outlined in Section 2. All the consumers in this economy share the same set of consumption functions and savings functions $\{g_t(s), h_t(s)\}_{t=0}^{T}$, where $s = (a, \widetilde{e}, \varepsilon)$ is a vector of individual state variables. The cross-sectional distributions of individual states are captured by the probability measures $\{\pi_t(\cdot)\}_{t=0}^{T}$, where each π_t is defined on the Borel algebra of S_t. Intuitively, $\pi_t(s)$ represents the share of age-t consumers whose current state is s. Since all consumers share the same level of initial asset a_0 and the same initial value $\widetilde{e}_0$, the probability measure $\pi_0(\cdot)$ is completely determined by the distribution $G_0(\cdot)$. The probability measure in all subsequent ages are defined recursively according to

$$\pi_{t+1}(B) = \int_{S_t} P_t(s, B)\, \pi_t(ds), \tag{9}$$

for any Borel set $B \subseteq S_{t+1}$. The stochastic kernel $P_t(s, B)$ is defined as

$$P_t(s, B) \equiv \Pr\{(\nu_{t+1}, \varepsilon_{t+1}) : (h_t(s), \widetilde{e}\nu_{t+1}, \varepsilon_{t+1}) \in B\}. \tag{10}$$

For each age group t, the economy-wide average level of wealth is determined by

$$\mathcal{W}_t \equiv \int_{S_t} h_t(s)\, \pi_t(ds). \tag{11}$$

The sequence $\{\mathcal{W}_t\}_{t=0}^{T}$ then forms the average life-cycle profile of wealth.

We are interested in how changes in the *riskiness* of the permanent and transitory shocks would affect the policy functions $\{g_t(s), h_t(s)\}_{t=0}^{T}$ and the average life-cycle profile of wealth $\{\mathcal{W}_t\}_{t=0}^{T}$ under a given set of prices. Throughout this section, we use the following criterion to compare the riskiness of two sets of distributions. Let $\mathbf{L}_1 \equiv \{L_{1,t}(\cdot)\}_{t=0}^{T}$ and $\mathbf{L}_2 \equiv \{L_{2,t}(\cdot)\}_{t=0}^{T}$ denote two sets of distribution functions for the permanent shocks $\{\nu_t\}$. These two sets of distribution functions are defined on the same compact interval $\Lambda \equiv [\underline{\nu}, \overline{\nu}]$. The distributions in $\mathbf{L}_1$ are said to be *more risky* than those in $\mathbf{L}_2$ if the inequality below holds for all $t \in \{0, 1, ..., T\}$ and for all concave function $f : \Lambda \to \mathbb{R}$,

$$\int_{\Lambda} f(\nu)\, dL_{1,t}(\nu) \leq \int_{\Lambda} f(\nu)\, dL_{2,t}(\nu),$$

provided that the integrals exist. As shown in Rothschild and Stiglitz (1970), this definition is equivalent to saying that each $L_{1,t}(\cdot)$ is a mean-preserving spread of $L_{2,t}(\cdot)$. It also means that the variance of $L_{1,t}(\cdot)$ is no less than that of $L_{2,t}(\cdot)$ in every period t. The same criterion is also used to compare the riskiness of two sets of distributions for the transitory shocks $\{\varepsilon_t\}$. Let $h_{j,t}(s)$ be the savings function in period t obtained under the distributions $\mathbf{L}_j$ (or $\mathbf{G}_j$), for $j \in \{1,2\}$. Lemma 2 summarizes the effects of greater income risk on individual savings. Specifically, it states that when the consumption functions are concave, an increase in the riskiness of the permanent shocks and purely transitory shocks would induce all consumers to accumulate more wealth.

Lemma 2 *Suppose the conditions in Theorem 1 are satisfied.*

(i) $\mathbf{L}_1$ *is more risky than* $\mathbf{L}_2$ *implies* $h_{1,t}(s) \geq h_{2,t}(s)$, *for all* $s \in S_t$ *and for all* $t \in \{0,1,...,T\}$.

(ii) $\mathbf{G}_1$ *is more risky than* $\mathbf{G}_2$ *implies* $h_{1,t}(s) \geq h_{2,t}(s)$, *for all* $s \in S_t$ *and for all* $t \in \{0,1,...,T\}$.

The main results of this section are presented in Theorem 2. It states that, when the consumption functions are concave, a mean-preserving increase in the permanent shocks and purely transitory shocks would raise the expected value of any increasing convex transformation $\Gamma(\cdot)$ of the savings function. This result can be obtained because $h_t(a, \widetilde{e}, \varepsilon)$ is convex in $(a, \widetilde{e})$ and convexity is preserved by any increasing convex transformation. Since $\Gamma(x) = x$ is increasing convex, it follows immediately that an increase in the riskiness of these shocks would raise the average level of wealth at every stage of the life cycle.

Theorem 2 *Suppose the conditions in Theorem 1 are satisfied.*

(i) $\mathbf{L}_1$ *is more risky than* $\mathbf{L}_2$ *implies*

$$\int_{S_t} \Gamma\left[h_{1,t}(s)\right] \pi_{1,t}(ds) \geq \int_{S_t} \Gamma\left[h_{2,t}(s)\right] \pi_{2,t}(ds),$$

for any continuous, increasing and convex function $\Gamma : A_{t+1} \to \mathbb{R}$, *and for all* $t \in \{0,1,...,T\}$.

(ii) $\mathbf{G}_1$ *is more risky than* $\mathbf{G}_2$ *implies*

$$\int_{S_t} \Gamma\left[h_{1,t}(s)\right] \pi_{1,t}(ds) \geq \int_{S_t} \Gamma\left[h_{2,t}(s)\right] \pi_{2,t}(ds),$$

for any continuous, increasing and convex function $\Gamma : A_{t+1} \rightarrow \mathbb{R}$, *and for all* $t \in \{0, 1, ..., T\}$.

The above results have two other implications regarding saving behavior in the presence of multiple risks. First, an anticipated increase in income risk in *any* future time period (both the distant and near-term future) can generate precautionary savings in period t. Second, when consumption functions are concave, the introduction of an independent risk in some future time periods will amplify the precautionary saving motive in period t. To illustrate these points, consider three sets of distribution functions for the purely transitory shocks, $\{\mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3\}$.[25] The distributions in $\mathbf{G}_2$ are identical to those in $\mathbf{G}_1$ except for one period, in particular $G_{2,t+1}(\cdot)$ is a mean-preserving spread of $G_{1,t+1}(\cdot)$. Similarly, $\mathbf{G}_3$ is identical to $\mathbf{G}_2$ except for period T, and $G_{3,T}(\cdot)$ is a mean-preserving spread of $G_{2,T}(\cdot)$. Let $h_{j,t}(s)$ be the savings function and $\mathcal{W}_{j,t}$ be the average level of wealth in period t under $\mathbf{G}_j$, for $j \in \{1, 2, 3\}$. Then our Lemma 2 and Theorem 2 imply the following: $h_{1,t}(s) \leq h_{2,t}(s) \leq h_{3,t}(s)$ for all $s \in S_t$, and $\mathcal{W}_{1,t} \leq \mathcal{W}_{2,t} \leq \mathcal{W}_{3,t}$. The interpretation of these rankings is as follows. When consumers decide how much to save in period t, they take into account not only the income shock in the near-term future (ε_{t+1}), but also those in the more far-off future, i.e., $\{\varepsilon_{t+2}, ..., \varepsilon_T\}$. When the consumption functions are concave, an increase in the riskiness of ε_{t+1} will increase both individual and aggregate savings, so that $h_{1,t}(s) \leq h_{2,t}(s)$ and $\mathcal{W}_{1,t} \leq \mathcal{W}_{2,t}$. Likewise, an anticipated increase in the riskiness of ε_T, or any other component of $\{\varepsilon_{t+2}, ..., \varepsilon_T\}$, will also encourage savings in period t, so that $h_{2,t}(s) \leq h_{3,t}(s)$ and $\mathcal{W}_{2,t} \leq \mathcal{W}_{3,t}$. Intuitively, one can think of $\{\varepsilon_{t+2}, ..., \varepsilon_T\}$ as some "background" risks that are independent of ε_{t+1}. Thus, our results imply that when the consumption functions are concave, an increase in these background risks will reinforce the precautionary saving motive in period t.

5 Concluding Remarks

The existence and consequences of precautionary savings have long been a subject of economic research. Earlier studies that analyzed a two-period model have produced important insights. Generalizing this analysis to a multi-period framework, however, proved to be very difficult. Partly for this reason, most of the existing studies have resorted to using numerical methods to investigate precautionary saving behavior under a particular analytical form of utility function. The present study takes a different

[25] The same argument can also be applied to permanent shocks.

path. Specifically, we analyze the theoretical foundations of precautionary saving behavior in a life-cycle model without limiting our attention to a particular class of utility functions. We are able to derive a set of novel conditions under which aggregate precautionary savings will occur. We believe the methodology in this paper, which built upon the one in Huggett (2004), can also be applied to models with other interesting features, such as endogenous labor supply, progressive taxation, and recursive preferences. We leave these to future research.

Appendix

Proof of Theorem 1

In the following proof, we focus on part (i) of the theorem. We will explain why the result in part (i) implies the result in part (ii) but not vice versa. The main ideas of the proof are as follows. For any $\varepsilon \in \Xi$ and $t \in \{0, 1, ..., T\}$, the function $g_t(a, \widetilde{e}, \varepsilon)$ is concave in $(a, \widetilde{e})$ if and only if its hypograph,

$$\mathcal{H}_t(\varepsilon) \equiv \{(c, a, \widetilde{e}) \in \mathcal{D} \times A_t \times \Delta_t : c \leq g_t(a, \widetilde{e}, \varepsilon)\},$$

is a convex set. The first step of the proof is to derive an alternate but equivalent expression for $\mathcal{H}_t(\varepsilon)$.[26] This alternate expression is favored because of its tractability. For each $(a, \widetilde{e}, \varepsilon) \in S_t$, define the constraint set

$$\mathcal{B}_t(a, \widetilde{e}, \varepsilon) \equiv \left\{c : \underline{c} \leq c \leq x(a, \widetilde{e}, \varepsilon) + \underline{a}_{t+1}\right\}.$$

For each $\varepsilon \in \Xi$, define a set $\mathcal{G}_t(\varepsilon)$ such that (i) $\mathcal{G}_t(\varepsilon)$ is a subset of $\mathcal{D} \times A_t \times \Delta_t$, and (ii) any $(c, a, \widetilde{e})$ in $\mathcal{G}_t(\varepsilon)$ satisfies $c \in \mathcal{B}_t(a, \widetilde{e}, \varepsilon)$ and

$$\phi(c) \geq \beta(1+r) \int_{\Xi} \int_{\Lambda} \phi\left[g_{t+1}\left(x(a, \widetilde{e}, \varepsilon) - c, \widetilde{e}\nu', \varepsilon'\right)\right] dL_{t+1}\left(\nu'\right) dG_{t+1}\left(\varepsilon'\right), \tag{12}$$

where $\phi(\cdot)$ is the marginal utility function, i.e., $\phi(c) = u'(c)$. We now show that $\mathcal{H}_t(\varepsilon)$ and $\mathcal{G}_t(\varepsilon)$ are equivalent. Fix $\varepsilon \in \Xi$. For any $(c, a, \widetilde{e}) \in \mathcal{H}_t(\varepsilon)$, it must be the case that $c \in \mathcal{B}_t(a, \widetilde{e}, \varepsilon)$ and $x(a, \widetilde{e}, \varepsilon) - c \geq x(a, \widetilde{e}, \varepsilon) - g_t(a, \widetilde{e}, \varepsilon)$. As stated in Lemma 1, the partial derivative of $V_{t+1}(a', \widetilde{e}', \varepsilon')$ with respect to a' is given by $(1+r)\phi[g_{t+1}(a', \widetilde{e}', \varepsilon')]$. Since $V_{t+1}(a', \widetilde{e}', \varepsilon')$ is strictly concave in a', it follows that $\phi[g_{t+1}(a', \widetilde{e}', \varepsilon')]$ is strictly decreasing in a'. Thus, we have

$$\phi\left[g_{t+1}\left(x(a, \widetilde{e}, \varepsilon) - g_t(a, \widetilde{e}, \varepsilon), \widetilde{e}\nu', \varepsilon'\right)\right] \geq \phi\left[g_{t+1}\left(x(a, \widetilde{e}, \varepsilon) - c, \widetilde{e}\nu', \varepsilon'\right)\right], \tag{13}$$

[26] The same step is also used in the proof of Lemma 1 in Huggett (2004). It is, however, necessary to include the details in here due to the differences in model specifications between the two work.

for all $(\nu', \varepsilon') \in \Lambda \times \Xi$. It follows that for any $(c, a, \widetilde{e}) \in \mathcal{H}_t(\varepsilon)$, we have

$$\begin{aligned}
\phi(c) &\geq \phi[g_t(a, \widetilde{e}, \varepsilon)] \\
&\geq \beta(1+r) \int_\Xi \int_\Lambda \phi\left[g_{t+1}\left(x(a, \widetilde{e}, \varepsilon) - g_t(a, \widetilde{e}, \varepsilon), \widetilde{e}\nu', \varepsilon'\right)\right] dL_{t+1}(\nu') \, dG_{t+1}(\varepsilon') \\
&\geq \beta(1+r) \int_\Xi \int_\Lambda \phi\left[g_{t+1}\left(x(a, \widetilde{e}, \varepsilon) - c, \widetilde{e}\nu', \varepsilon'\right)\right] dL_{t+1}(\nu') \, dG_{t+1}(\varepsilon').
\end{aligned}$$

The second inequality uses the Euler equation and the third inequality follows from (13). This shows that $\mathcal{H}_t(\varepsilon) \subseteq \mathcal{G}_t(\varepsilon)$. Next, pick any $(c, a, \widetilde{e})$ in $\mathcal{G}_t(\varepsilon)$ and suppose the contrary that $c > g_t(a, \widetilde{e}, \varepsilon)$. If $g_t(a, \widetilde{e}, \varepsilon) = x(a, \widetilde{e}, \varepsilon) + \underline{a}_{t+1}$, then any feasible consumption must be no greater than $g_t(a, \widetilde{e}, \varepsilon)$ and hence there is a contradiction. Consider the case when $x(a, \widetilde{e}, \varepsilon) + \underline{a}_{t+1} \geq c > g_t(a, \widetilde{e}, \varepsilon)$. This has two implications: (i) $h_t(a, \widetilde{e}, \varepsilon) > -\underline{a}_{t+1}$, and (ii) $h_t(a, \widetilde{e}, \varepsilon) > x(a, \widetilde{e}, \varepsilon) - c$. The first inequality implies that the Euler equation holds with equality under $g_t(a, \widetilde{e}, \varepsilon)$. Thus, we have

$$\begin{aligned}
\phi(c) &< \phi[g_t(a, \widetilde{e}, \varepsilon)] = \beta(1+r) \int_\Xi \int_\Lambda \phi\left[g_{t+1}\left(h_t(a, \widetilde{e}, \varepsilon), \widetilde{e}\nu', \varepsilon'\right)\right] dL_{t+1}(\nu') \, dG_{t+1}(\varepsilon') \\
&< \beta(1+r) \int_\Xi \int_\Lambda \phi\left[g_{t+1}\left(x(a, \widetilde{e}, \varepsilon) - c, \widetilde{e}\nu', \varepsilon'\right)\right] dL_{t+1}(\nu') \, dG_{t+1}(\varepsilon').
\end{aligned}$$

This means $(c, a, \widetilde{e}) \notin \mathcal{G}_t(\varepsilon)$ which gives rise to a contradiction. Hence, $\mathcal{G}_t(\varepsilon) \subseteq \mathcal{H}_t(\varepsilon)$. This establishes the equivalence between $\mathcal{H}_t(\varepsilon)$ and $\mathcal{G}_t(\varepsilon)$.

Since $\phi(\cdot)$ is strictly decreasing, the inequality in (12) is equivalent to

$$c \leq \phi^{-1}\left\{\beta(1+r) \int_\Xi \int_\Lambda \phi\left[g_{t+1}\left(x(a, \widetilde{e}, \varepsilon) - c, \widetilde{e}\nu', \varepsilon'\right)\right] dL_{t+1}(\nu') \, dG_{t+1}(\varepsilon')\right\}.$$

Define a function $\Psi_{t+1} : A_{t+1} \times \Delta_t \to \mathcal{D}$ according to

$$\Psi_{t+1}(a', \widetilde{e}) \equiv \phi^{-1}\left\{\beta(1+r) \int_\Xi \int_\Lambda \phi\left[g_{t+1}\left(a', \widetilde{e}\nu', \varepsilon'\right)\right] dL_{t+1}(\nu') \, dG_{t+1}(\varepsilon')\right\}. \tag{14}$$

Then the set $\mathcal{H}_t(\varepsilon)$ can be rewritten as

$$\mathcal{H}_t(\varepsilon) \equiv \{(c, a, \widetilde{e}) \in \mathcal{D} \times A_t \times \Delta_t : c \in \mathcal{B}_t(a, \widetilde{e}, \varepsilon) \text{ and } c \leq \Psi_{t+1}(x(a, \widetilde{e}, \varepsilon) - c, \widetilde{e})\}.$$

This set is convex if $\Psi_{t+1}(a', \widetilde{e})$ is jointly concave in $(a', \widetilde{e})$. To see this, pick any $(c_1, a_1, \widetilde{e}_1)$ and

$(c_2, a_2, \widetilde{e}_2)$ in $\mathcal{H}_t(\varepsilon)$. Define $c_\delta \equiv \delta c_1 + (1-\delta) c_2$ for any $\delta \in (0,1)$. Similarly define a_δ and $\widetilde{e}_\delta$. Since $\mathcal{D} \times A_t \times \Delta_t$ is a convex set, we have $(c_\delta, a_\delta, \widetilde{e}_\delta) \in \mathcal{D} \times A_t \times \Delta_t$. Also, we have $c_\delta \in \mathcal{B}_t(a_\delta, \widetilde{e}_\delta, \varepsilon)$. If $\Psi_{t+1}(\cdot)$ is concave, then

$$\begin{aligned} \Psi_{t+1}(x(a_\delta, \widetilde{e}_\delta, \varepsilon) - c_\delta, \widetilde{e}_\delta) &\geq \delta \Psi_{t+1}(x(a_1, \widetilde{e}_1, \varepsilon) - c_1, \widetilde{e}_1) + (1-\delta)\Psi_{t+1}(x(a_2, \widetilde{e}_2, \varepsilon) - c_2, \widetilde{e}_2) \\ &\geq \delta c_1 + (1-\delta) c_2 \equiv c_\delta. \end{aligned}$$

This means $(c_\delta, a_\delta, \widetilde{e}_\delta) \in \mathcal{H}_t(\varepsilon)$. Hence, if $\Psi_{t+1}(\cdot)$ is a concave function, then $g_t(a, \widetilde{e}, \varepsilon)$ is concave in $(a, \widetilde{e})$. The converse, however, is not necessarily true.

To establish the result in part (ii), we first define the hypograph of $g_t(a, \widetilde{e}, \varepsilon)$ for a given $\widetilde{e} \in \Delta_t$. Using the same procedure, we can derive an alternate expression for this hypograph, which involves the same function $\Psi_{t+1}(a', \widetilde{e})$ as defined in (14). If $\Psi_{t+1}(a', \widetilde{e})$ is concave in a' for each given $\widetilde{e} \in \Delta_t$, then the consumption function is concave in (a, ε). Note that concavity of $\Psi_{t+1}(\cdot)$ implies concavity of $\Psi_{t+1}(\cdot, \widetilde{e})$ for a given $\widetilde{e}$, but the converse is not true in general. Hence, concavity in $(a, \widetilde{e})$ implies concavity in (a, ε), but not vice versa.

Case 1: Quadratic Utility

Suppose $u'''(c) = 0$ for all $c \in \mathcal{D}$. Then the marginal utility function can be expressed as $\phi(c) = \vartheta_1 + \vartheta_2 c$, with $\vartheta_2 < 0$ and $\vartheta_1 + \vartheta_2 \underline{c} > 0$. It follows that

$$\Psi_{t+1}(a', \widetilde{e}) \equiv \frac{[\beta(1+r) - 1]\vartheta_1}{\vartheta_2} + \beta(1+r) \int_\Xi \int_\Lambda g_{t+1}(a', \widetilde{e}\nu', \varepsilon') \, dL_{t+1}(\nu') \, dG_{t+1}(\varepsilon').$$

Concavity of $\Psi_{t+1}(\cdot)$ follows immediately from an inductive argument. In the terminal period, the policy function is $g_T(a, \widetilde{e}, \varepsilon) \equiv w\widetilde{e}\varepsilon + (1+r)a$, which is linear in $(a, \widetilde{e})$ for all $\varepsilon \in \Xi$. Suppose $g_{t+1}(a', \widetilde{e}', \varepsilon')$ is concave in $(a', \widetilde{e}')$ for any given $\varepsilon' \in \Xi$. Since concavity is preserved under integration, it follows that $\Psi_{t+1}(a', \widetilde{e})$ is also concave in $(a', \widetilde{e})$. Hence $\mathcal{H}_t(\varepsilon)$ is a convex set and $g_t(a, \widetilde{e}, \varepsilon)$ is concave in $(a, \widetilde{e})$ for all $\varepsilon \in \Xi$.

Case 2: Utility with Strictly Positive Third Derivative

Suppose now $u'''(c) > 0$ for all $c \in \mathcal{D}$. Again we will use an inductive argument to establish the concavity of $\Psi_{t+1}(\cdot)$. Suppose $g_{t+1}(a', \widetilde{e}', \varepsilon')$ is concave in $(a', \widetilde{e}')$ for all $\varepsilon' \in \Xi$ and for some $t+1 \leq T$.

We first establish the concavity of $\Psi_{t+1}(\cdot)$ for the case when both $G_{t+1}(\cdot)$ and $L_{t+1}(\cdot)$ are discrete distributions defined on some finite point sets. We then extend this result to continuous distributions.

Suppose $L_{t+1}(\cdot)$ is a discrete distribution with positive masses over a set of real numbers $\{\overline{\nu}_1, ..., \overline{\nu}_J\}$, with $\overline{\nu}_j \in \Lambda$ for all j. Similarly, suppose $G_{t+1}(\cdot)$ is a discrete distribution with positive masses over a set of real numbers $\{\overline{\varepsilon}_1, ..., \overline{\varepsilon}_K\}$, with $\overline{\varepsilon}_k \in \Xi$ for all k. Both J and K are finite. Define $\mathcal{P}_{t+1}(j,k) \equiv \Pr\{(\nu_{t+1}, \varepsilon_{t+1}) = (\overline{\nu}_j, \overline{\varepsilon}_k)\}$ for each pair $(\overline{\nu}_j, \overline{\varepsilon}_k)$. The function $\Psi_{t+1}(\cdot)$ defined in (14) can be rewritten as

$$\Psi_{t+1}(a', \widetilde{e}) \equiv \phi^{-1}\left\{\beta(1+r)\sum_{j,k}\mathcal{P}_{t+1}(j,k)\,\phi\left[g_{t+1}(a', \widetilde{e}\overline{\nu}_j, \overline{\varepsilon}_k)\right]\right\}. \tag{15}$$

Define a related function $\theta : (\underline{c}, \infty)^N \rightarrow \mathcal{D}$ according to

$$\theta(\mathbf{y}) \equiv \phi^{-1}\left\{\beta(1+r)\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\,\phi(y_n)\right\}, \tag{16}$$

We now establish the concavity of $\Psi_{t+1}(\cdot)$ by proving a set of claims.

Claim 1: $\Psi_{t+1}(\cdot)$ *is concave if* $\theta(\cdot)$ *is concave.*

Proof of Claim 1 Pick any $(a'_1, \widetilde{e}_1)$ and $(a'_2, \widetilde{e}_2)$ in $A_{t+1} \times \Delta_t$. Define $a'_\delta \equiv \delta a'_1 + (1-\delta)a'_2$ for any $\delta \in (0,1)$. Similarly define $\widetilde{e}_\delta$. Since $g_{t+1}(a', \widetilde{e}', \varepsilon')$ is concave in $(a', \widetilde{e}')$ and $\phi(\cdot)$ is strictly decreasing, we have

$$\phi\left[g_{t+1}(a'_\delta, \widetilde{e}_\delta\overline{\nu}_j, \overline{\varepsilon}_k)\right] \leq \phi\left[\delta g_{t+1}(a'_1, \widetilde{e}_1\overline{\nu}_j, \overline{\varepsilon}_k) + (1-\delta)\,g_{t+1}(a'_2, \widetilde{e}_2\overline{\nu}_j, \overline{\varepsilon}_k)\right],$$

for all possible $(\overline{\nu}_j, \overline{\varepsilon}_k)$. Taking the expectation over all possible $(\overline{\nu}_j, \overline{\varepsilon}_k)$ gives

$$\begin{aligned} & \beta(1+r)\sum_{j,k}\mathcal{P}_{t+1}(j,k)\,\phi\left[g_{t+1}(a'_\delta, \widetilde{e}_\delta\overline{\nu}_j, \overline{\varepsilon}_k)\right] \\ \leq \; & \beta(1+r)\sum_{j,k}\mathcal{P}_{t+1}(j,k)\,\phi\left[\delta g_{t+1}(a'_1, \widetilde{e}_1\overline{\nu}_j, \overline{\varepsilon}_k) + (1-\delta)\,g_{t+1}(a'_2, \widetilde{e}_2\overline{\nu}_j, \overline{\varepsilon}_k)\right]. \end{aligned}$$

Since the inverse function $\phi^{-1}(\cdot)$ is also strictly decreasing, we can write

$$\begin{aligned}
\Psi_{t+1}\left(a'_\delta, \widetilde{e}_\delta\right) &\equiv \phi^{-1}\left\{\beta(1+r)\sum_{j,k}\mathcal{P}_{t+1}(j,k)\,\phi\left[g_{t+1}\left(a'_\delta, \widetilde{e}_\delta\overline{\nu}_j, \overline{\varepsilon}_k\right)\right]\right\} \\
&\geq \phi^{-1}\left\{\beta(1+r)\sum_{j,k}\mathcal{P}_{t+1}(j,k)\,\phi\left[\delta g_{t+1}\left(a'_1, \widetilde{e}_1\overline{\nu}_j, \overline{\varepsilon}_k\right) + (1-\delta)\,g_{t+1}\left(a'_2, \widetilde{e}_2\overline{\nu}_j, \overline{\varepsilon}_k\right)\right]\right\}.
\end{aligned}$$

To express this more succinctly, define an index $n \equiv (j-1)\times K + k$ for any pair (j,k). Set $N \equiv J \times K$. Using the index n, we can reduce the double summation to a single one. Define two sets of positive real numbers $\mathbf{x} \equiv \{x_n\}_{n=1}^N$ and $\mathbf{y} \equiv \{y_n\}_{n=1}^N$ according to $x_n \equiv g_{t+1}(a'_1, \widetilde{e}_1\overline{\nu}_j, \overline{\varepsilon}_k)$ and $y_n \equiv g_{t+1}(a'_2, \widetilde{e}_2\overline{\nu}_j, \overline{\varepsilon}_k)$ for all (j,k). With a slight abuse of notation, we will use $\mathcal{P}_{t+1}(n)$ to replace $\mathcal{P}_{t+1}(j,k)$. Then the above inequality can be expressed as

$$\Psi_{t+1}\left(a'_\delta, \widetilde{e}_\delta\right) \geq \phi^{-1}\left\{\beta(1+r)\sum_{n=1}^N \mathcal{P}_{t+1}(n)\,\phi\left[\delta x_n + (1-\delta)\,y_n\right]\right\} = \theta\left(\delta\mathbf{x} + (1-\delta)\,\mathbf{y}\right).$$

The equality follows from (16). Concavity of $\theta(\cdot)$ implies

$$\begin{aligned}
\theta\left(\delta\mathbf{x} + (1-\delta)\,\mathbf{y}\right) &\geq \delta\theta(\mathbf{x}) + (1-\delta)\,\theta(\mathbf{y}) \\
&\equiv \delta\phi^{-1}\left\{\beta(1+r)\sum_{n=1}^N \mathcal{P}_{t+1}(n)\,\phi(x_n)\right\} + (1-\delta)\,\phi^{-1}\left\{\beta(1+r)\sum_{n=1}^N \mathcal{P}_{t+1}(n)\,\phi(y_n)\right\} \\
&= \delta\Psi_{t+1}\left(a'_1, \widetilde{e}_1\right) + (1-\delta)\,\Psi_{t+1}\left(a'_2, \widetilde{e}_2\right).
\end{aligned}$$

Hence, $\Psi_{t+1}(\cdot)$ is concave if $\theta(\cdot)$ is concave. This establishes Claim 1.

Before we state our next claim, we need to introduce some additional notions. Define $\eta : \mathcal{D} \to \mathbb{R}$ according to

$$\eta(c) = \frac{\left[\phi'(c)\right]^2}{\phi''(c)} \equiv \frac{\left[u''(c)\right]^2}{u'''(c)}, \tag{17}$$

and $\Phi : \mathbb{R}_+ \to \mathbb{R}$ according to

$$\Phi(m) \equiv \eta\left[\phi^{-1}(m)\right]. \tag{18}$$

Both $\eta(\cdot)$ and $\Phi(\cdot)$ are strictly positive if $u'''(\cdot)$ is strictly positive. The following condition provides a useful intermediate step to characterize the concavity of $\theta(\cdot)$.

Condition A For any arbitrary discrete probability measure with masses $(\mu_1, ..., \mu_N)$ on a set of points $(\psi_1, ..., \psi_N) \in \mathbb{R}_+^N$, the function $\Phi(\cdot)$ defined in (18) satisfies the following condition

$$\Phi\left[\beta(1+r)\sum_{n=1}^{N}\mu_n\psi_n\right] \geq \beta(1+r)\sum_{n=1}^{N}\mu_n\Phi(\psi_n). \tag{19}$$

Claim 2: *The function $\theta(\cdot)$ defined in (16) is concave if Condition A is satisfied.*

Proof of Claim 2 To establish the concavity of $\theta(\cdot)$, we follow the approach as in Hardy, Littlewood and Pólya (1952) p.85-88.[27] The function $\theta(\mathbf{y})$ is concave if and only if its Hessian matrix is negative semi-definite. Let $\mathbf{H}(\mathbf{y}) = [\mathbf{h}_{m,n}(\mathbf{y})]$ be the Hessian matrix of $\theta(\cdot)$ evaluated at a point $\mathbf{y}$. This matrix is negative semi-definite if and only if $\boldsymbol{\varpi}^T \cdot \mathbf{H}(\mathbf{y})\boldsymbol{\varpi} \leq 0$ for any column vector $\boldsymbol{\varpi} \in \mathbb{R}^N$. The elements of the Hessian matrix can be derived as follows. First, rewrite (16) as

$$\phi[\theta(\mathbf{y})] = \beta(1+r)\sum_{n=1}^{N}\mathcal{P}_t(n)\phi(y_n).$$

Differentiating this with respect to y_n gives

$$\phi'[\theta(\mathbf{y})]\mathbf{h}_n(\mathbf{y}) = \beta(1+r)\mathcal{P}_t(n)\phi'(y_n), \tag{20}$$

$$\Rightarrow \mathbf{h}_n(\mathbf{y}) = \beta(1+r)\mathcal{P}_t(n)\frac{\phi'(y_n)}{\phi'[\theta(\mathbf{y})]}, \tag{21}$$

where $\mathbf{h}_n(\mathbf{y}) \equiv \partial\theta(\mathbf{y})/\partial y_n$. Differentiating (20) with respect to y_m gives

$$\phi''[\theta(\mathbf{y})]\mathbf{h}_m(\mathbf{y})\mathbf{h}_n(\mathbf{y}) + \phi'[\theta(\mathbf{y})]\mathbf{h}_{m,n}(\mathbf{y}) = 0, \tag{22}$$

if $m \neq n$, and

$$\phi''[\theta(\mathbf{y})][\mathbf{h}_n(\mathbf{y})]^2 + \phi'[\theta(\mathbf{y})]\mathbf{h}_{n,n}(\mathbf{y}) = \beta(1+r)\mathcal{P}_t(n)\phi''(y_n), \tag{23}$$

[27] For the case when $\beta(1+r) = 1$, we can use the same line of argument as in Ben-Tal and Teboulle (1986) to show that $\theta(\cdot)$ is concave if and only if $I(c) \equiv -u''(c)/u'''(c)$ is concave. Their proof, however, cannot be extended to the case when $\beta(1+r) < 1$.

if $m = n$. Combining (21) and (22) gives

$$\begin{aligned}\mathbf{h}_{m,n}(\mathbf{y}) &= -\frac{\phi''[\theta(\mathbf{y})]}{\phi'[\theta(\mathbf{y})]}\mathbf{h}_m(\mathbf{y})\,\mathbf{h}_n(\mathbf{y}) \\ &= -[\beta(1+r)]^2\,\mathcal{P}_t(m)\,\mathcal{P}_t(n)\,\phi'(y_m)\,\phi'(y_n)\,\frac{\phi''[\theta(\mathbf{y})]}{\{\phi'[\theta(\mathbf{y})]\}^3},\end{aligned}$$

for $m \neq n$. Similarly, combining (21) and (23) gives

$$\mathbf{h}_{n,n}(\mathbf{y}) = \beta(1+r)\,\mathcal{P}_t(n)\,\frac{\phi''(y_n)}{\phi'[\theta(\mathbf{y})]} - \left[\beta(1+r)\,\mathcal{P}_t(n)\,\phi'(y_n)\right]^2\frac{\phi''[\theta(\mathbf{y})]}{\{\phi'[\theta(\mathbf{y})]\}^3}.$$

For any $\boldsymbol{\varpi} \in \mathbb{R}^N$, we have

$$\begin{aligned}&\boldsymbol{\varpi}^T \cdot \mathbf{H}(\mathbf{y})\,\boldsymbol{\varpi} \\ =\;& \beta(1+r)\,\frac{\sum_{n=1}^N \mathcal{P}_t(n)\,\varpi_n^2\phi''(y_n)}{\phi'[\theta(\mathbf{y})]} - [\beta(1+r)]^2\left[\sum_{n=1}^N \mathcal{P}_t(n)\,\phi'(y_n)\right]^2\frac{\phi''[\theta(\mathbf{y})]}{\{\phi'[\theta(\mathbf{y})]\}^3} \\ =\;& \beta(1+r)\,\frac{\phi''[\theta(\mathbf{y})]}{\{\phi'[\theta(\mathbf{y})]\}^3}\left[\sum_{n=1}^N \mathcal{P}_t(n)\,\varpi_n^2\phi''(y_n)\right]\left\{\frac{\{\phi'[\theta(\mathbf{y})]\}^2}{\phi''[\theta(\mathbf{y})]} - \beta(1+r)\,\frac{\left[\sum_{n=1}^N \mathcal{P}_t(n)\,\varpi_n\phi'(y_n)\right]^2}{\left[\sum_{n=1}^N \mathcal{P}_t(n)\,\varpi_n^2\phi''(y_n)\right]}\right\}.\end{aligned}$$

Since $\beta(1+r) > 0$ and $\phi(\cdot)$ is strictly decreasing and strictly convex, it follows that $\boldsymbol{\varpi}^T \cdot \mathbf{H}(\mathbf{y})\,\boldsymbol{\varpi} \leq 0$ if and only if

$$\frac{\{\phi'[\theta(\mathbf{y})]\}^2}{\phi''[\theta(\mathbf{y})]} \geq \beta(1+r)\,\frac{\left[\sum_{n=1}^N \mathcal{P}_{t+1}(n)\,\varpi_n\phi'(y_n)\right]^2}{\left[\sum_{n=1}^N \mathcal{P}_{t+1}(n)\,\varpi_n^2\phi''(y_n)\right]}, \qquad \text{for all } \boldsymbol{\varpi} \in \mathbb{R}^N. \tag{24}$$

Using the definitions in (17) and (18), we can rewrite the left-hand side of (24) as

$$\begin{aligned}\frac{\{\phi'[\theta(\mathbf{y})]\}^2}{\phi''[\theta(\mathbf{y})]} &\equiv \eta[\theta(\mathbf{y})] = \eta\left[\phi^{-1}\left\{\beta(1+r)\sum_{n=1}^N \mathcal{P}_{t+1}(n)\,\phi(y_n)\right\}\right] \\ &= \boldsymbol{\Phi}\left[\beta(1+r)\sum_{n=1}^N \mathcal{P}_{t+1}(n)\,\phi(y_n)\right].\end{aligned}$$

Using Condition A, we can obtain

$$\frac{\{\phi'[\theta(\mathbf{y})]\}^2}{\phi''[\theta(\mathbf{y})]} = \Phi\left[\beta(1+r)\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\phi(y_n)\right]$$
$$\geq \beta(1+r)\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\Phi[\phi(y_n)] = \beta(1+r)\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\frac{[\phi'(y_n)]^2}{\phi''(y_n)}. \qquad (25)$$

The final step is to show that (25) implies (24). Define two sets of real numbers $\{\mathbf{b}_n\}_{n=1}^N$ and $\{\mathbf{d}_n\}_{n=1}^N$ according to $\mathbf{b}_n \equiv [\mathcal{P}_{t+1}(n)\phi''(y_n)]^{\frac{1}{2}}\varpi_n$ and $\mathbf{d}_n \equiv \left\{\mathcal{P}_{t+1}(n)[\phi'(y_n)]^2/\phi''(y_n)\right\}^{\frac{1}{2}}$. By the Cauchy-Schwartz inequality,

$$\left(\sum_{n=1}^{N}\mathbf{b}_n^2\right)\left(\sum_{n=1}^{N}\mathbf{d}_n^2\right) = \left[\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\varpi_n^2\phi''(y_n)\right]\left[\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\frac{[\phi'(y_n)]^2}{\phi''(y_n)}\right]$$
$$\geq \left(\sum_{n=1}^{N}\mathbf{b}_n\mathbf{d}_n\right)^2$$
$$= \left[\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\varpi_n\phi'(y_n)\right]^2.$$

Since $\phi''(\cdot) > 0$, this yields

$$\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\frac{[\phi'(y_n)]^2}{\phi''(y_n)} \geq \frac{\left[\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\varpi_n\phi'(y_n)\right]^2}{\left[\sum_{n=1}^{N}\mathcal{P}_{t+1}(n)\varpi_n^2\phi''(y_n)\right]},$$

for any $\boldsymbol{\varpi} \in \mathbb{R}^N$. Hence (25) implies (24) and this establishes Claim 2.

Claim 3: *Condition A is satisfied if $\Phi(\cdot)$ is a concave function.*

Proof of Claim 3 First, note that when $\beta(1+r) = 1$, the inequality in (19) is identical to Jensen's inequality. Hence, Condition A is satisfied by any concave function $\Phi(\cdot)$ when $\beta(1+r) = 1$. Consider the case when $\beta(1+r) \in (0,1)$. Since $u'''(c) > 0$ for all $c \in \mathcal{D}$, we have $\Phi(m) \geq 0$ for all $m \geq 0$. This, together with the concavity of $\Phi(\cdot)$, implies $\Phi(\delta m) \geq \delta\Phi(m)$ for all $m \geq 0$ and $\delta \in [0,1]$. Let m_1 and m_2 be two arbitrary real numbers that satisfy $m_1 > m_2 > 0$. Then there exists $\delta \in (0,1)$ such that $m_2 = \delta m_1$. It follows that

$$\Phi(m_2) = \Phi(\delta m_1) \geq \delta\Phi(m_1)$$

$$\Rightarrow \frac{\Phi(m_2)}{m_2} \geq \frac{\Phi(m_1)}{m_1},$$

i.e., $m^{-1}\Phi(m)$ is a nonincreasing function in $m > 0$.

Consider an arbitrary discrete probability distribution with masses $(\mu_1, ..., \mu_N)$ on a set of points $(\psi_1, ..., \psi_N) \in \mathbb{R}_+^N$. First consider the case when $\sum_{n=1}^N \mu_n \psi_n = 0$. Since $\psi_i \geq 0$ for all i, this can happen when either (i) $\psi_n = 0$ for all n, or (ii) $\psi_n > 0$ for some n but $\mu_n = 0$. In both cases, we have

$$\Phi\left[\beta(1+r)\sum_{n=1}^N \mu_n \psi_n\right] = \Phi(0) > \beta(1+r)\sum_{n=1}^N \mu_n \Phi(\psi_n) = \beta(1+r)\Phi(0),$$

as $\beta(1+r) \in (0,1)$. Hence, Condition A is satisfied. Next, consider the case when $\sum_{n=1}^N \mu_n \psi_n > 0$. Since $m^{-1}\Phi(m)$ is a nonincreasing function in $m > 0$, we have

$$\frac{\Phi\left[\beta(1+r)\sum_{n=1}^N \mu_n \psi_n\right]}{\beta(1+r)\sum_{n=1}^N \mu_n \psi_n} \geq \frac{\Phi\left[\sum_{n=1}^N \mu_n \psi_n\right]}{\sum_{n=1}^N \mu_n \psi_n} \geq \frac{\sum_{n=1}^N \mu_n \Phi(\psi_n)}{\sum_{n=1}^N \mu_n \psi_n}.$$

The second inequality follows from the concavity of $\Phi(\cdot)$. The desired result can be obtained by rearranging terms. This establishes Claim 3.

Claim 4: *The function $\Phi(\cdot)$ is concave if and only if the inverse of absolute prudence $I(\cdot)$ is concave.*

Proof of Claim 4 Since $u(\cdot)$ is four-times differentiable, the first-order derivative of $I(\cdot)$ and $\Phi(\cdot)$ exist at each point in $\mathcal{D}$. Let $u^{(4)}(\cdot)$ denote the fourth-order derivative of the utility function. Then the derivative of $I(c)$ is given by

$$I'(c) = -1 + \frac{u''(c)\, u^{(4)}(c)}{\left[u'''(c)\right]^2}.$$

The above expression is non-increasing over $\mathcal{D}$ if and only if $I(\cdot)$ is a concave function. The derivative of $\Phi(\cdot)$, on the other hand, is given by

$$\Phi'(m) = \frac{\eta'\left[\phi^{-1}(m)\right]}{\phi'\left[\phi^{-1}(m)\right]},$$

where

$$\eta'(c) = 2\phi'(c) - \frac{\left[\phi'(c)\right]^2 \phi'''(c)}{\left[\phi''(c)\right]^2}.$$

Combining these two expressions gives

$$\begin{aligned}\Phi'(m) &= 2-\frac{\phi'\left[\phi^{-1}(m)\right]\phi'''\left[\phi^{-1}(m)\right]}{\left\{\phi''\left[\phi^{-1}(m)\right]\right\}^2}\\ &= 2-\frac{u''\left[\phi^{-1}(m)\right]u^{(4)}\left[\phi^{-1}(m)\right]}{\left\{u'''\left[\phi^{-1}(m)\right]\right\}^2}=1-I'\left[\phi^{-1}(m)\right].\end{aligned}$$

Since $\phi^{-1}(\cdot)$ is strictly decreasing, this means $\Phi'(\cdot)$ is non-increasing if and only if $I'(\cdot)$ is non-increasing. This establishes Claim 4.

Through Claim 1 to Claim 4, we have shown that if the inverse of absolute prudence $I(\cdot)$ is a concave function, then the $\Psi_{t+1}(\cdot)$ in (15) is concave. This in turn implies that the hypograph of $g_t(a,\widetilde{e},\varepsilon)$ is a convex set for each fixed $\varepsilon\in\Xi$. This proves the desired result for the case when both $G_{t+1}(\cdot)$ and $L_{t+1}(\cdot)$ are discrete distributions defined on some finite point sets.

Suppose now both $G_{t+1}(\cdot)$ and $L_{t+1}(\cdot)$ are continuous distributions defined on the compact intervals $\Xi\equiv[\underline{\varepsilon},\overline{\varepsilon}]$ and $\Lambda\equiv[\underline{\nu},\overline{\nu}]$, respectively. Fix $(a',\widetilde{e})\in A_{t+1}\times\Delta_t$. Let J and K be two positive integers. Let $\{\overline{\nu}_0,...,\overline{\nu}_J\}$ be an arbitrary partition of Λ so that $\underline{\nu}=\overline{\nu}_0\leq...\leq\overline{\nu}_J=\overline{\nu}$. Define a set of real numbers $\{p_1,...,p_J\}$ according to $p_j\equiv L_{t+1}(\overline{\nu}_j)-L_{t+1}(\overline{\nu}_{j-1})$ for each $j\geq1$, and a step function

$$\widetilde{L}_J(\nu')\equiv\sum_{j=1}^{J}\chi_j(\nu')\,L_{t+1}(\overline{\nu}_{j-1}),$$

where $\chi_j(\nu')$ equals one if $\nu'\in[\overline{\nu}_{j-1},\overline{\nu}_j]$ and zero otherwise. This step function converges pointwise to $L_{t+1}(\cdot)$ when J is sufficiently large. Similarly, let $\{\overline{\varepsilon}_0,...,\overline{\varepsilon}_K\}$ be an arbitrary partition of Ξ so that $\underline{\varepsilon}=\overline{\varepsilon}_0\leq...\leq\overline{\varepsilon}_K=\overline{\varepsilon}$. Define a set of positive real numbers $\{q_1,...,q_K\}$ so that $q_k\equiv G_{t+1}(\overline{\varepsilon}_k)-G_{t+1}(\overline{\varepsilon}_{k-1})$ for each $k\geq1$. Define the step function

$$\widetilde{G}_K(\varepsilon')\equiv\sum_{k=1}^{K}\widetilde{\chi}_k(\varepsilon')\,G_{t+1}(\overline{\varepsilon}_{k-1}),$$

where $\widetilde{\chi}_k(\varepsilon')$ equals one if $\varepsilon'\in[\overline{\varepsilon}_{k-1},\overline{\varepsilon}_k]$ and zero otherwise. This step function converges pointwise to $G_{t+1}(\cdot)$ when K is sufficiently large. These conditions are sufficient to ensure that

$$\sum_{j,k}p_jq_k\phi\left[g_{t+1}\left(a',\widetilde{e}\overline{\nu}_j,\overline{\varepsilon}_k\right)\right]\to\int_{\Xi}\int_{\Lambda}\phi\left[g_{t+1}\left(a',\widetilde{e}\nu',\varepsilon'\right)\right]dL_{t+1}(\nu')\,dG_{t+1}(\varepsilon'),$$

for any given $(a', \widetilde{e}) \in A_{t+1} \times \Delta_t$, when J and K are sufficiently large. Set $N = J \times K$ and define a function $\Psi^N_{t+1}(a', \widetilde{e})$ according to

$$\Psi^N_{t+1}\left(a', \widetilde{e}\right) \equiv \phi^{-1}\left\{\beta\left(1+r\right)\sum_{j,k} p_j q_k \phi\left[g_{t+1}\left(a', \widetilde{e}\overline{\nu}_j, \overline{\varepsilon}_k\right)\right]\right\}.$$

Our earlier result shows that $\Psi^N_{t+1}(a', \widetilde{e})$ is jointly concave in its arguments for any positive integer N. By the continuity of $\phi^{-1}(\cdot)$, $\Psi^N_{t+1}(\cdot)$ converges to the function in (14) pointwise. Hence, $\left\{\Psi^N_{t+1}(\cdot)\right\}$ forms a sequence of finite concave function on $A_{t+1} \times \Delta_t$ that converges pointwise to $\Psi_{t+1}(\cdot)$. By Theorem 10.8 in Rockafellar (1970), the limiting function $\Psi_{t+1}(\cdot)$ is also a concave function on $A_{t+1} \times \Delta_t$. This completes the proof of Theorem 1.

Proof of Lemma 2

The proof of part (ii) is largely similar to that of part (i), thus we only present the proof of part (i) here. Start from age $T-1$. Suppose the contrary that $h_{2,T-1}(a, \widetilde{e}, \varepsilon) > h_{1,T-1}(a, \widetilde{e}, \varepsilon) \geq -\underline{a}_T$, for some $(a, \widetilde{e}, \varepsilon) \in S_{T-1}$. Then for any $(\nu', \varepsilon') \in \Lambda \times \Xi$, we have

$$w\widetilde{e}\nu'\varepsilon' + (1+r)\, h_{2,T-1}(a, \widetilde{e}, \varepsilon) > w\widetilde{e}\nu'\varepsilon' + (1+r)\, h_{1,T-1}(a, \widetilde{e}, \varepsilon).$$

Since $\phi(\cdot)$ is strictly decreasing, this means

$$\phi\left[w\widetilde{e}\nu'\varepsilon' + (1+r)\, h_{2,T-1}(a, \widetilde{e}, \varepsilon)\right] < \phi\left[w\widetilde{e}\nu'\varepsilon' + (1+r)\, h_{1,T-1}(a, \widetilde{e}, \varepsilon)\right], \tag{26}$$

for all (ν', ε'). Since $\phi(\cdot)$ is also strictly convex, the expression $\phi\left[w\widetilde{e}\nu'\varepsilon' + (1+r)\, h_{2,T-1}(a, \widetilde{e}, \varepsilon)\right]$ is strictly convex in ν' when $(a, \widetilde{e}, \varepsilon, \varepsilon')$ are held fixed. This, together with the assumption that $L_{1,T}(\cdot)$ is more risky than $L_{2,T}(\cdot)$, implies

$$\int_\Lambda \phi\left[w\widetilde{e}\nu'\varepsilon' + (1+r)\, h_{2,T-1}(a, \widetilde{e}, \varepsilon)\right] dL_{1,T}\left(\nu'\right) > \int_\Lambda \phi\left[w\widetilde{e}\nu'\varepsilon' + (1+r)\, h_{2,T-1}(a, \widetilde{e}, \varepsilon)\right] dL_{2,T}\left(\nu'\right), \tag{27}$$

for any given $\varepsilon' \in \Xi$. Using the Euler equation, we can obtain

$$\begin{aligned}
\phi\left[g_{1,T-1}(a,\widetilde{e},\varepsilon)\right] &\geq \beta(1+r)\int_{\Xi}\int_{\Lambda}\phi\left[w\widetilde{e}\nu'\varepsilon' + (1+r)h_{1,T-1}(a,\widetilde{e},\varepsilon)\right]dL_{1,T}(\nu')\,dG_T(\varepsilon') \\
&> \beta(1+r)\int_{\Xi}\int_{\Lambda}\phi\left[w\widetilde{e}\nu'\varepsilon' + (1+r)h_{2,T-1}(a,\widetilde{e},\varepsilon)\right]dL_{1,T}(\nu')\,dG_T(\varepsilon') \\
&\geq \beta(1+r)\int_{\Xi}\int_{\Lambda}\phi\left[w\widetilde{e}\nu'\varepsilon' + (1+r)h_{2,T-1}(a,\widetilde{e},\varepsilon)\right]dL_{2,T}(\nu')\,dG_T(\varepsilon') \\
&= \phi\left[g_{2,T-1}(a,\widetilde{e},\varepsilon)\right].
\end{aligned}$$

The second line uses (26) while the third line uses (27). The last line follows from the assumption that $h_{2,T-1}(a,\widetilde{e},\varepsilon) > -\underline{a}_T$. The above result implies $g_{2,T-1}(a,\widetilde{e},\varepsilon) > g_{1,T-1}(a,\widetilde{e},\varepsilon)$ which contradicts $h_{2,T-1}(a,\widetilde{e},\varepsilon) > h_{1,T-1}(a,\widetilde{e},\varepsilon)$. Hence, $h_{2,T-1}(a,\widetilde{e},\varepsilon) \leq h_{1,T-1}(a,\widetilde{e},\varepsilon)$ for all $(a,\widetilde{e},\varepsilon) \in S_{T-1}$.

Suppose $h_{2,t+1}(a',\widetilde{e}',\varepsilon') \leq h_{1,t+1}(a',\widetilde{e}',\varepsilon')$ for all $(a',\widetilde{e}',\varepsilon') \in S_{t+1}$ and for $t+1 \leq T-1$. This means $g_{2,t+1}(a',\widetilde{e}',\varepsilon') \geq g_{1,t+1}(a',\widetilde{e}',\varepsilon')$ for all $(a',\widetilde{e}',\varepsilon') \in S_{t+1}$. Suppose the contrary that $h_{2,t}(a,\widetilde{e},\varepsilon) > h_{1,t}(a,\widetilde{e},\varepsilon) \geq -\underline{a}_{t+1}$, for some $(a,\widetilde{e},\varepsilon) \in S_t$. Then for any $(\nu',\varepsilon') \in \Lambda \times \Xi$, we have

$$\begin{aligned}
g_{1,t+1}\left(h_{1,t}(a,\widetilde{e},\varepsilon), \widetilde{e}\nu', \varepsilon'\right) &\leq g_{2,t+1}\left(h_{1,t}(a,\widetilde{e},\varepsilon), \widetilde{e}\nu', \varepsilon'\right) \\
&< g_{2,t+1}\left(h_{2,t}(a,\widetilde{e},\varepsilon), \widetilde{e}\nu', \varepsilon'\right).
\end{aligned}$$

The first inequality follows from the induction hypothesis. The second inequality uses the fact that $g_{2,t+1}(a',\widetilde{e}',\varepsilon')$ is strictly increasing in a'. Since $\phi(\cdot)$ is strictly decreasing, this means

$$\phi\left[g_{2,t+1}\left(h_{2,t}(a,\widetilde{e},\varepsilon), \widetilde{e}\nu', \varepsilon'\right)\right] < \phi\left[g_{1,t+1}\left(h_{1,t}(a,\widetilde{e},\varepsilon), \widetilde{e}\nu', \varepsilon'\right)\right],$$

which is analogous to (26). If we can show that $\phi\left[g_{2,t+1}\left(h_{2,t}(a,\widetilde{e},\varepsilon), \widetilde{e}\nu', \varepsilon'\right)\right]$ is convex in ν', then a contradiction can be obtained by using the same argument. Pick any ν_1' and ν_2' in Λ. For any $\delta \in (0,1)$, define $\nu_\delta' \equiv \delta\nu_1' + (1-\delta)\nu_2'$. Since $g_{2,t+1}(a',\widetilde{e}',\varepsilon')$ is concave in $\widetilde{e}'$, we can obtain

$$\begin{aligned}
& g_{2,t+1}\left(h_{2,t}(a,\widetilde{e},\varepsilon), \widetilde{e}\nu_\delta', \varepsilon'\right) \\
\geq\; & \delta g_{2,t+1}\left(h_{2,t}(a,\widetilde{e},\varepsilon), \widetilde{e}\nu_1', \varepsilon'\right) + (1-\delta)\, g_{2,t+1}\left(h_{2,t}(a,\widetilde{e},\varepsilon), \widetilde{e}\nu_2', \varepsilon'\right).
\end{aligned}$$

Since $\phi(\cdot)$ is strictly decreasing and strictly convex, this means

$$\begin{aligned} & \phi\left[g_{2,t+1}\left(h_{2,t}\left(a,\widetilde{e},\varepsilon\right),\widetilde{e}\nu'_{\delta},\varepsilon'\right)\right] \\ \leq\; & \phi\left[\delta g_{2,t+1}\left(h_{2,t}\left(a,\widetilde{e},\varepsilon\right),\widetilde{e}\nu'_{1},\varepsilon'\right)+\left(1-\delta\right)g_{2,t+1}\left(h_{2,t}\left(a,\widetilde{e},\varepsilon\right),\widetilde{e}\nu'_{2},\varepsilon'\right)\right] \\ \leq\; & \delta\phi\left[g_{2,t+1}\left(h_{2,t}\left(a,\widetilde{e},\varepsilon\right),\widetilde{e}\nu'_{1},\varepsilon'\right)\right]+\left(1-\delta\right)\phi\left[g_{2,t+1}\left(h_{2,t}\left(a,\widetilde{e},\varepsilon\right),\widetilde{e}\nu'_{2},\varepsilon'\right)\right]. \end{aligned}$$

We can now use the same argument as in the terminal period to show that $g_{2,t}(a,\widetilde{e},\varepsilon) > g_{1,t}(a,\widetilde{e},\varepsilon)$ which is inconsistent with $h_{2,t}(a,\widetilde{e},\varepsilon) > h_{1,t}(a,\widetilde{e},\varepsilon)$. Hence, $h_{2,t}(a,\widetilde{e},\varepsilon) \leq h_{1,t}(a,\widetilde{e},\varepsilon)$ for all $(a,\widetilde{e},\varepsilon) \in S_t$. This proves part (i) of the lemma.

Proof of Theorem 2

Again, the proof of part (ii) parallels that of part (i), hence we only present the proof of part (i). Unlike Huggett (2004), which uses the Markov operator to establish his results, we use the sequential approach in the following proof. The sequential approach is useful because it allows us to fully exploit the assumption that $\{\varepsilon_0,\varepsilon_1,...,\varepsilon_t,\nu_1,...,\nu_t\}$ is a set of independent random variables. The proof is divided into a number of steps.

Step 1 First, we derive an alternate expression for the expectation of $\Gamma[h_{j,t}(s_t)]$ using the distributions $\{G_1,...,G_t,L_{j,1},...,L_{j,t}\}$. By Theorem 8.3 of Stokey, Lucas and Prescott (1989), we can obtain

$$\int_{S_t}\Gamma\left[h_{j,t}\left(s_t\right)\right]\pi_{j,t}\left(ds_t\right)=\int_{S_{t-1}}\left[\int_{S_t}\Gamma\left[h_{j,t}\left(s_t\right)\right]P_{j,t-1}\left(s_{t-1},ds_t\right)\right]\pi_{j,t-1}\left(ds_{t-1}\right),$$

where

$$\int_{S_t}\Gamma\left[h_{j,t}\left(s_t\right)\right]P_{j,t-1}\left(s_{t-1},ds_t\right)=\int_{\Xi}\int_{\Lambda}\Gamma\left[h_{j,t}\left(h_{j,t-1}\left(s_{t-1}\right),\widetilde{e}_{t-1}\nu_t,\varepsilon_t\right)\right]dL_{j,t}\left(\nu_t\right)dG_t\left(\varepsilon_t\right)\equiv F_{j,t}\left(s_{t-1}\right),$$

for all $s_{t-1}\in S_{t-1}$. Applying the same theorem, we can obtain

$$\int_{S_{t-1}}F_{j,t}\left(s_{t-1}\right)\pi_{j,t-1}\left(ds_{t-1}\right)=\int_{S_{t-2}}\left[\int_{S_{t-1}}F_{j,t}\left(s_{t-1}\right)P_{j,t-2}\left(s_{t-2},ds_{t-1}\right)\right]\pi_{j,t-2}\left(ds_{t-2}\right),$$

where

$$\int_{S_{t-1}} F_{j,t}(s_{t-1}) P_{j,t-2}(s_{t-2}, ds_{t-1}) = \int_{\Xi}\int_{\Lambda} F_{j,t}\left[h_{j,t-2}(s_{t-2}), \widetilde{e}_{t-2}\nu_{t-1}, \varepsilon_{t-1}\right] dL_{j,t-1}(\nu_{t-1})\, dG_{t-1}(\varepsilon_{t-1}),$$

for all $s_{t-2} \in S_{t-2}$. By repeating the same procedure, we can obtain

$$\begin{aligned} & \int_{S_t} \Gamma\left[h_{j,t}(s_t)\right] \pi_{j,t}(ds_t) \qquad (28) \\ = & \int_{\Xi} \cdots \int_{\Lambda} \Gamma\left[h_{j,t}(\cdots h_{j,1}(h_{j,0}(a_0, \widetilde{e}_0, \varepsilon_0), \widetilde{e}_0\nu_1, \varepsilon_1) \cdots, \widetilde{e}_{t-1}\nu_t, \varepsilon_t)\right] dL_{j,t}(\nu_t) \ldots dL_{j,1}(\nu_1)\, dG_t(\varepsilon_t) \ldots dG_0(\varepsilon_0), \end{aligned}$$

where $\widetilde{e}_{t-1}\nu_t = \widetilde{e}_0\nu_1 \ldots \nu_t$.

Let $\boldsymbol{\varepsilon}^t = \{\varepsilon_0, ..., \varepsilon_t\}$ denote a history of transitory shocks up to age t, and $\boldsymbol{\nu}^t = \{\nu_1, ..., \nu_t\}$ denote a history of permanent shocks. For $j \in \{1, 2\}$ and for $t \in \{0, 1, ..., T\}$, define a function $f_{j,t}$ according to

$$f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right) \equiv h_{j,t}(\cdots h_{j,1}(h_{j,0}(a_0, \widetilde{e}_0, \varepsilon_0), \widetilde{e}_0\nu_1, \varepsilon_1) \cdots, \widetilde{e}_{t-1}\nu_t, \varepsilon_t).$$

Hence, we can write

$$\int_{S_t} \Gamma\left[h_{j,t}(s_t)\right] \pi_{j,t}(ds_t) \equiv \int_{\Xi} \cdots \int_{\Lambda} \Gamma\left[f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] dL_{j,t}(\nu_t) \ldots dL_{j,1}(\nu_1)\, dG_t(\varepsilon_t) \ldots dG_0(\varepsilon_0).$$

Step 2 We now show that $f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)$ is convex in every single ν_τ, $\tau \leq t$, when all other arguments $(\varepsilon_0, ..., \varepsilon_t, \nu_1, ..., \nu_{\tau-1}, \nu_{\tau+1}, ...\nu_t)$ are fixed. An induction argument is used to establish this result. When $t = 1$, we have

$$f_{j,1}(\varepsilon_0, \varepsilon_1, \nu_1) \equiv h_{j,1}(h_{j,0}(a_0, \widetilde{e}_0, \varepsilon_0), \widetilde{e}_0\nu_1, \varepsilon_1).$$

By Theorem 1, $h_{j,1}(\cdot)$ is convex in its second argument when the other arguments are held constant. Thus, $f_{j,1}(\varepsilon_0, \varepsilon_1, \nu_1)$ is convex in ν_1 for any $(\varepsilon_0, \varepsilon_1)$. Suppose the desired result is true for $f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)$. In the next period,

$$f_{j,t+1}\left(\boldsymbol{\varepsilon}^{t+1}, \boldsymbol{\nu}^{t+1}\right) \equiv h_{j,t+1}\left(f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right), \widetilde{e}_0\nu_1 \ldots \nu_{t+1}, \varepsilon_{t+1}\right).$$

Since $h_{j,t+1}(\cdot)$ is convex in its second argument, $f_{j,t+1}\left(\boldsymbol{\varepsilon}^{t+1}, \boldsymbol{\nu}^{t+1}\right)$ is convex in ν_{t+1} when all other arguments are held constant. Fix $\tau \leq t$. Pick any $\nu_{1,\tau}$ and $\nu_{2,\tau}$ from Λ. For any $\delta \in (0, 1)$, define

$\nu_{\delta,\tau} \equiv \delta\nu_{1,\tau} + (1-\delta)\,\nu_{2,\tau}$. Define two histories of permanent shocks which differ only in terms of ν_τ, i.e., $\boldsymbol{\nu}_i^t \equiv \{\nu_1, ..., \nu_{i,\tau}, ...\nu_t\}$ for $i \in \{1,2\}$. Define $\boldsymbol{\nu}_\delta^t \equiv \delta\boldsymbol{\nu}_1^t + (1-\delta)\,\boldsymbol{\nu}_2^t$. Similarly, define $\widetilde{e}_{i,t} \equiv \widetilde{e}_0 \times \nu_1 \times \ldots \times \nu_{i,\tau} \times \ldots \times \nu_t$ for $i \in \{1,2\}$. Then we have $\widetilde{e}_{\delta,t} = \widetilde{e}_0 \times \nu_1 \times \ldots \times \nu_{\delta,\tau} \times \ldots \times \nu_t = \delta\widetilde{e}_{1,t} + (1-\delta)\,\widetilde{e}_{2,t}$. By the induction hypothesis, we have

$$f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}_\delta^t\right) \leq \delta f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}_1^t\right) + (1-\delta)\, f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}_2^t\right).$$

Since $h_{j,t+1}(\cdot)$ is increasing in its first argument and joint convex in its first two arguments, we have

$$\begin{aligned}
& h_{j,t+1}\left(f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}_\delta^t\right), \widetilde{e}_{\delta,t}\nu_{t+1}, \varepsilon_{t+1}\right) \\
\leq\ & h_{j,t+1}\left(\delta f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}_1^t\right) + (1-\delta)\, f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}_2^t\right), \widetilde{e}_{\delta,t}\nu_{t+1}, \varepsilon_{t+1}\right) \\
\leq\ & \delta h_{j,t+1}\left(f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}_1^t\right), \widetilde{e}_{1,t}\nu_{t+1}, \varepsilon_{t+1}\right) + (1-\delta)\, h_{j,t+1}\left(f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}_2^t\right), \widetilde{e}_{2,t}\nu_{t+1}, \varepsilon_{t+1}\right).
\end{aligned}$$

This establishes the convexity of $f_{j,t+1}\left(\boldsymbol{\varepsilon}^{t+1}, \boldsymbol{\nu}^{t+1}\right)$ in every single ν_τ, for $\tau \leq t+1$. Since convexity is preserved by any increasing convex transformation, it follows that $\Gamma\left[f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right]$ is also convex in every single ν_τ, $\tau \leq t$, when all other arguments $(\varepsilon_0, ..., \varepsilon_t, \nu_1, ..., \nu_{\tau-1}, \nu_{\tau+1}, ...\nu_t)$ are held constant.

The above argument, however, is not valid if we assume that $\ln\widetilde{e}_t$ follows a stationary AR(1) process, i.e.,

$$\ln\widetilde{e}_t = \rho\ln\widetilde{e}_{t-1} + \nu_t,$$

where $\rho \in (0,1)$ and ν_t is a serially independent random variable. To see this, it suffice to consider $f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)$ for $t \in \{1,2\}$, which are now given by

$$f_{j,1}(\varepsilon_0, \varepsilon_1, \nu_1) = h_{j,1}\left(h_{j,0}\left(a_0, \widetilde{e}_0^\rho, \varepsilon_0\right), \widetilde{e}_0^\rho\nu_1, \varepsilon_1\right),$$

$$f_{j,2}(\varepsilon_0, \varepsilon_1, \varepsilon_2, \nu_1, \nu_2) = h_{j,2}\left(f_{j,1}(\varepsilon_0, \varepsilon_1, \nu_1), \widetilde{e}_0^\rho\nu_1^\rho\nu_2, \varepsilon_2\right).$$

The main problem is that $\widetilde{e}_0^\rho\nu_1^\rho\nu_2$ is *strictly concave* in ν_1 for any $\rho \in (0,1)$. Thus, even if $h_{j,2}(a, \widetilde{e}, \varepsilon)$ is increasing and convex in $(a, \widetilde{e})$, there is no guarantee that the composite function $f_{j,2}(\cdot)$ is convex in ν_1. This problem, however, will not affect the proof of part (ii). Specifically, if $h_{j,t}(a, \widetilde{e}, \varepsilon)$ is increasing and convex in (a, ε), then $f_{j,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)$ is convex in *every single* element in $\boldsymbol{\varepsilon}^t = \{\varepsilon_1, ..., \varepsilon_t\}$, regardless of the specification of the persistent shock.

Step 3 We now show that $f_{1,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right) \geq f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)$ for any possible history $\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)$ and for all t. Fix $\left(\boldsymbol{\varepsilon}^T, \boldsymbol{\nu}^T\right)$. By Lemma 2, we have $h_{1,0}\left(a_0, \widetilde{e}_0, \varepsilon_0\right) \geq h_{2,0}\left(a_0, \widetilde{e}_0, \varepsilon_0\right)$. When $t = 1$, we have

$$\begin{aligned} f_{1,1}\left(\varepsilon_0, \varepsilon_1, \nu_1\right) &\equiv h_{1,1}(h_{1,0}\left(a_0, \widetilde{e}_0, \varepsilon_0\right), \widetilde{e}_0\nu_1, \varepsilon_1) \\ &\geq h_{2,1}(h_{1,0}\left(a_0, \widetilde{e}_0, \varepsilon_0\right), \widetilde{e}_0\nu_1, \varepsilon_1) \\ &\geq h_{2,1}(h_{2,0}\left(a_0, \widetilde{e}_0, \varepsilon_0\right), \widetilde{e}_0\nu_1, \varepsilon_1) \equiv f_{2,1}\left(\varepsilon_0, \varepsilon_1, \nu_1\right). \end{aligned}$$

The second line again uses the result in Lemma 2. The third line uses the fact that $h_{2,1}(\cdot)$ is increasing in its first argument. Suppose $f_{1,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right) \geq f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)$ for some $t \geq 1$. Using the same line of argument, we can obtain

$$\begin{aligned} f_{1,t+1}\left(\boldsymbol{\varepsilon}^{t+1}, \boldsymbol{\nu}^{t+1}\right) &\equiv h_{1,t+1}\left(f_{1,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right), \widetilde{e}_t\nu_{t+1}, \varepsilon_{t+1}\right) \\ &\geq h_{2,t+1}\left(f_{1,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right), \widetilde{e}_t\nu_{t+1}, \varepsilon_{t+1}\right) \\ &\geq h_{2,t+1}\left(f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right), \widetilde{e}_t\nu_{t+1}, \varepsilon_{t+1}\right) \equiv f_{2,t+1}\left(\boldsymbol{\varepsilon}^{t+1}, \boldsymbol{\nu}^{t+1}\right). \end{aligned}$$

This establishes the desired result. Since $\Gamma(\cdot)$ is an increasing function, we have $\Gamma\left[f_{1,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] \geq \Gamma\left[f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right]$ for any possible history $\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)$ and for all t.

Step 4 Finally, we will show that the expected value of $\Gamma\left[f_{1,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] \equiv \Gamma\left[h_{1,t}\left(s_t\right)\right]$ is no less than the expected value of $\Gamma\left[f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] \equiv \Gamma\left[h_{2,t}\left(s_t\right)\right]$. Fix $\boldsymbol{\varepsilon}^t$. Then we have

$$\begin{aligned} & \int_\Lambda \dots \int_\Lambda \Gamma\left[f_{1,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] dL_{1,t}\left(\nu_t\right) dL_{1,t-1}\left(\nu_{t-1}\right) \dots dL_{1,1}\left(\nu_1\right) \\ \geq & \int_\Lambda \dots \int_\Lambda \Gamma\left[f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] dL_{1,t}\left(\nu_t\right) dL_{1,t-1}\left(\nu_{t-1}\right) \dots dL_{1,1}\left(\nu_1\right) \\ \geq & \int_\Lambda \dots \int_\Lambda \Gamma\left[f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] dL_{2,t}\left(\nu_t\right) dL_{1,t-1}\left(\nu_{t-1}\right) \dots dL_{1,1}\left(\nu_1\right) \\ \geq & \int_\Lambda \dots \int_\Lambda \Gamma\left[f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] dL_{2,t}\left(\nu_t\right) dL_{2,t-1}\left(\nu_{t-1}\right) \dots dL_{1,1}\left(\nu_1\right) \\ & \dots \\ \geq & \int_\Lambda \dots \int_\Lambda \Gamma\left[f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] dL_{2,t}\left(\nu_t\right) dL_{2,t-1}\left(\nu_{t-1}\right) \dots dL_{2,1}\left(\nu_1\right). \end{aligned}$$

The first inequality follows from the result in step 3. The second inequality uses the following facts: (i) $\Gamma\left[f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right]$ is convex in ν_t when all other components are held constant, and (ii) $L_{1,t}(\cdot)$ is more risky than $L_{2,t}(\cdot)$. Since convexity is preserved by integration, it follows that $\int_{\Lambda} \Gamma\left[f_{2,t}\left(\boldsymbol{\varepsilon}^t, \boldsymbol{\nu}^t\right)\right] dL_{2,t}(\nu_t)$ is convex in ν_{t-1} when $(\nu_1, ..., \nu_{t-2})$ are fixed. The third inequality follows from this property. The last line can be obtained by repeating the same argument for all preceding periods. This procedure is valid because $\{\nu_1, ..., \nu_t\}$ is a set of independent random variables. Since this ordering is true for any given history $\boldsymbol{\varepsilon}^t$, the desired result follows by taking the expectation over all possible $\boldsymbol{\varepsilon}^t$. This proves part (i) of the theorem.

References

[1] ABOWD, J. and CARD, D. (1989), "On the Covariance Structure of Earnings and Hours Change," *Econometrica*, **57**, 411-445.

[2] AIYAGARI, R. (1994), "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, **109**, 659-684.

[3] BEN-TAL, A. and TEBOULLE, M. (1986), "Expected Utility, Penalty Functions and Duality in Stochastic Nonlinear Programming", *Management Science*, **32**, 1445–1466

[4] BLUNDELL, R. and STOKER, T. M. (1999), "Consumption and the Timing of Income Risk," *European Economic Review*, **43**, 475-507.

[5] BROWNING, M. and LUSARDI, A. (1996), "Household Saving: Micro Theories and Micro Facts," *Journal of Economic Literature*, **34**, 1797-1855.

[6] CARROLL, C. D. (1992), "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, **2**, 61-156.

[7] CARROLL, C. D. (1997), "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, **112**, 1-55.

[8] CARROLL, C. D. (2011), "Theoretical Foundations of Buffer Stock Saving," (Manuscript)

[9] CARROLL, C. D. and KIMBALL, M. S. (1996), "On the Concavity of the Consumption Function," *Econometrica*, **64**, 981-992.

[10] CARROLL, C. D. and KIMBALL, M. S. (2005), "Liquidity Constraints and Precautionary Saving," (Manuscript)

[11] CARROLL, C. D. and SAMWICK, A. A. (1997), "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, **40**, 41-71.

[12] CARROLL, C. D. and SAMWICK, A. A. (1998), "How Important is Precautionary Saving?" *Review of Economics and Statistics*, **80**, 410-419.

[13] DEATON, A. (1991), "Saving and Liquidity Constraints," *Econometrica*, **59**, 1221-1248.

[14] GHIGLINO, C. (2005) "Wealth Inequality and Dynamic Stability," *Journal of Economic Theory*, **124**, 106-115.

[15] GHIGLINO, C., VENDITTI, A. (2011) "Wealth Distribution and Output Fluctuations," *Journal of Economic Theory*, **146**, 2478-2509.

[16] GOLLIER, C. (2001a), "Wealth Inequality and Asset Pricing," *Review of Economics Studies*, **68**, 181-203.

[17] GOLLIER, C. (2001b), *The Economics of Risk and Time,* (Cambridge: MIT Press).

[18] GOLLIER, C. and ZECKHAUSER, R. J. (2002), "Horizon Length and Portfolio Risk," *Journal of Risk and Uncertainty*, **24**, 195-212.

[19] GOURINCHAS, P.-O. and PARKER, J. A. (2002), "Consumption over the Life Cycle," *Econometrica*, **70**, 47-89.

[20] HARDY, G., LITTLEWOOD, J. E. and PÓLYA, G. (1952), *Inequalities*, 2nd Edition (Cambridge: Cambridge University Press).

[21] HEALTHCOTE, J., STORESLETTEN, K. and VIOLANTE, G. L. (2009), "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*, **1**, 319-354.

[22] HUGGETT, M. (2004), "Precautionary Wealth Accumulation", *Review of Economics Studies*, **71**, 769-781.

[23] HUGGETT, M. and OSPINA S. (2001), "Aggregate Precautionary Savings: When is the Third Derivative Irrelevant?" *Journal of Monetary Economics*, **48**, 373-396.

[24] HUGGETT, M. and VIDON, E. (2002), "Precautionary Wealth Accumulation: A Positive Third Derivative is not Enough," *Economics Letters*, **76**, 323-329.

[25] JAPPELLI, T. and PISTAFERRI, L. (2014), "Fiscal Policy and MPC Heterogeneity," *American Economic Journal: Macroeconomics,* **6**, 107-136.

[26] JOHNSON, D. S., PARKER, J. A., and SOULELES, N. S. (2006), "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, **96**, 1589-1610.

[27] KIMBALL, M. S. (1990), "Precautionary Saving in the Small and in the Large," *Econometrica*, **58**, 53-73.

[28] LELAND, H. E. (1968), "Saving and Uncertainty: The Precautionary Demand for Saving," *Quarterly Journal of Economics*, **82**, 465-473.

[29] LUDVIGSON, S. (1999), "Consumption and Credit: A Model of Time-Varying Liquidity Constraints," *Review of Economics and Statistics*, **81**, 434-447.

[30] LUSARDI, A. (1998), "On the Importance of the Precautionary Saving Motive," *American Economic Review* (*Papers and Proceedings*), **88**, 449-453.

[31] MAZZOCCO, M. (2004), "Saving, Risk Sharing, and Preferences for Risk," *American Economic Review*, **94**, 1169-1182.

[32] MENDELSON, H. and AMIHUD, Y. (1982), "Optimal Consumption Policy under Uncertain Income," *Management Science*, **28**, 683-697.

[33] MILLER, B. L. (1976), "The Effect on Optimal Consumption of Increased Uncertainty in Labor Income in the Multiperiod Case," *Journal of Economic Theory*, **13**, 154-167.

[34] MOFFITT, R. A. and GOTTSCHALK, P. (2011), "Trends in the Covariance Structure of Earnings in the U.S.: 1969-1987," *Journal of Economic Inequality*, **9**, 439-459.

[35] PARKER, J. A. and PRESTON, B. (2005), "Precautionary Saving and Consumption Fluctuations," *American Economic Review*, **95**, 1119-1143.

[36] RABAULT, G. (2002), "When do Borrowing Constraints Bind? Some New Results on the Income Fluctuation Problem," *Journal of Economic Dynamics and Control*, **26**, 217-245.

[37] ROCKAFELLAR, R. T. (1970), *Convex Analysis* (Princeton: Princeton University Press).

[38] ROTHSCHILD, M. and STIGLITZ, J. (1970), "Increasing Risk I: A Definition," *Journal of Economic Theory*, **2**, 225-243.

[39] SCHECHTMAN, J. (1976), "An Income Fluctuation Problem," *Journal of Economic Theory*, **12**, 218-241.

[40] SIBLEY, D. S. (1975), “Permanent and Transitory Income Effects in a Model of Optimal Consumption with Wage Income Uncertainty,” *Journal of Economic Theory*, **11**, 68-82.

[41] STOKEY, N. L., LUCAS, R. E. and PRESCOTT, E. C. (1989), *Recursive Methods in Economic Dynamics*, (Cambridge: Harvard University Press).

[42] STORESLETTEN, K., TELMER, C. I. and YARON, A. (2004a), “Consumption and Risk Sharing over the Life Cycle,” *Journal of Monetary Economics*, **51**, 609-633.

[43] STORESLETTEN, K., TELMER, C. I. and YARON, A. (2004b), “Cyclical Dynamics in Idiosyncratic Labor Market Risk,” *Journal of Political Economy*, **112**, 695-717.

[44] ZELDES, S. P. (1989), “Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence,” *Quarterly Journal of Economics*, **104**, 275-298.