Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Monetary Economics 51 (2004) 1645–1681
Journal of MONETARY ECONOMICS
www.elsevier.com/locate/econbase

Precautionary saving and partially observed income ☆

Neng Wang*

Graduate School of Business, Columbia University, 3022 Broadway, Uris Hall 812, New York, NY 10027, USA

Received 18 August 2003; received in revised form 3 May 2004; accepted 17 June 2004
Available online 11 November 2004

Abstract

I propose an intertemporal precautionary saving model in which the agent's labor income is subject to (possibly correlated) shocks with different degrees of persistence and volatility. However, he only observes his total income, not individual components. I show that partial observability of individual components of income gives rise to additional precautionary saving due to estimation risk, the error associated with estimating individual components of income. This additional precautionary saving is higher, when estimation risk is greater. Compared with a precautionary agent who is otherwise identical, but ignores estimation risk, the rational agent consumes less at the beginning of his life, but consumes more later, because of larger wealth accumulated from savings for estimation risk. The utility cost of ignoring estimation risk is also quantified in closed form.

JEL classification: D11; D91; E21

Keywords: Precautionary saving; Separation principle; Incomplete information; Kalman filter

☆I thank Steve Davis, Darrell Duffie, Jeremy Greenwood, Steve Grenadier, Fatih Guvenen, Bob Hall, Ken Judd, Yingcong Lan, Luigi Pistaferri, Charles Plosser, Tom Sargent, John Shoven, Ken Singleton, Chris Telmer, Steve Zeldes, and especially the anonymous referee, for their helpful comments. Yangchun Chu provided excellent research assistance.
*Tel.: + 1 212 854 3869.
E-mail address: nw2128@columbia.edu (N. Wang).

0304-3932/$ - see front matter
doi:10.1016/j.jmoneco.2004.06.001

1. Introduction

The idea that income is subject to both permanent and transitory shocks dates at least back to Friedman (1957) and Muth (1960). Recently, much empirical evidence supports the view that individual's labor income has components with different degrees of persistence and volatility.[1] However, when it comes to analyze individual's optimal intertemporal consumption, vast majority of the literature assumes a univariate labor-income process,[2] probably for technical convenience. While the univariate income process is a reasonable approximation for many issues, it nonetheless ignores, among other things, the effect of estimation risk on intertemporal optimal consumption rule. By estimation risk, I mean the uncertainty in estimating the individual components of income, based on only observing total income. Intuitively, a smaller variance of estimation error indicates a lower estimation risk.

Exceptions to consumption models with univariate income processes include Goodfriend (1992), Pischke (1995), Quah (1990) and a collection of papers in Hansen and Sargent (1991), among others. These studies assume that the individual agent's income is subject to both permanent and transitory shocks. Goodfriend (1992) postulates that the individual agent lacks contemporaneous information on aggregate shocks.[3] Pischke (1995) supposes that the agent only observes his total income, not individual components of income. Quah (1990) assumes that the agent observes both components of the income process, while the econometrician only has data on the economic agent's total income. Hansen and Sargent (1991) contain detailed treatments on estimation and inference of multi-variate linear time series.

While these studies differ in their informational structure assumptions, they all postulate that the individual agent's optimal consumption rule is given by the permanent-income hypothesis (PIH) of Friedman (1957). The PIH states that consumption is equal to the annuity value of "total wealth" given by the sum of financial wealth (cumulative savings) and "human wealth," the discounted expected value of future income, using the risk-free rate. This in turn implies that changes in individual agent's consumption are not predictable (Hall, 1978). However, the conditions under which the PIH rule is optimal are quite restrictive and

[1] See MaCurdy (1982), Hall and Mishkin (1982), Abowd and Card (1989).

[2] See Zeldes (1989), Deaton (1991), and Aiyagari (1994) for example. Carroll (1997) considers an income process with both permanent and transitory components. The transitory component is assumed to be white noise. Therefore, his problem is effectively a univariate problem. Storesletten et al. (2004) study consumption inequality in an equilibrium overlapping generations model, by postulating an income process that is given by the sum of the agent's fixed effect, a permanent component and a transitory component. As in other papers, they also assume that the agent is able to perfectly observe each individual components of his income process.

[3] Goodfriend (1992) shows that the martingale property of consumption will not hold at the aggregate level due to imperfect information aggregation, even though the martingale property of consumption holds by assumption at the individual's level.

un-realistic.[4] Moreover, imposing the PIH rule excludes any possible effect of estimation risk on consumption, because the PIH rule assumes away any precautionary motive, by assumption.

Unlike the existing literature, this paper acknowledges incomplete information about individual components of labor income, and incorporates precautionary saving motive into an intertemporal consumption model with partially observed income. Precautionary motive is one of the most fundamental reasons for people to save, a point mentioned repeatedly in Friedman (1957),[5] and first modeled by Leland (1968) in a two-period setting. Since the pathbreaking work of Hall (1978), precautionary saving has attracted enormous attention in the consumption literature.[6] Precautionary saving models in the literature have either postulated a univariate labor income process or simply abstracted away from the issue of estimation risk, by assuming that individual components of income are observable. However, the agent often does not have complete information about individual components of his income. In those cases, forecasting is necessary and inevitably imperfect. As a result, forecasting error shall naturally affect his saving decision. A rational precautionary agent (whose marginal utility is convex, (Kimball, 1990)) recognizes this estimation-induced uncertainty and thus increases his saving optimally for precaution. Intuitively, a higher variance of forecasting error implies a bigger precautionary saving demand. This paper confirms this intuition in an intertemporal setting, and quantifies the effect of estimation risk on consumption, wealth accumulation, and lifetime utility. In a closely related setting, Guvenen (2004) argues that the individual agent may be uncertain about his own income profile and thus faces a signal extraction problem. Guvenen (2004) finds empirically that individual income persistence is substantially lowered after taking the partial observability of the individual's income profile. His analysis provides empirical evidence on the relevance of estimation risk of the individual income process forthe agent's consumption saving decisions.

In order to show that estimation risk gives rises to additional precautionary saving in a simple and tractable model, I assume that (i) the agent has time-additive expected constant-absolute-risk-averse (CARA) utility as in Caballero (1991), and (ii) the joint dynamics of his unobservable bi-variate income process is a conditionally homoskedastic Gaussian process. The CARA-utility-based model is able to capture the precautionary motive in an analytically tractable way.[7] The

[4] To justify the PIH in an intertemporal framework, we need to assume both (i) quadratic utility, to turn off precautionary motives (Hall, 1978), and (ii) equality between the subjective discount rate and the interest rate, in order to rule out the dis-saving for lack of patience. Neither assumption is plausible. Motivated by this observation, Wang (2003) shows that the PIH rule may also be obtained in a general equilibrium (Bewley-style) model with *ex ante* identical, but *ex post* heterogenous agents. Each agent in that economy solves a Caballero-style precautionary saving model.

[5] It is interesting to note that Milton Friedman's most known contribution to the consumption literature is the PIH consumption rule, which exactly ignores precautionary saving motive.

[6] See Caballero (1990, 1991), Zeldes (1989), Deaton (1991), Carroll (1997), and Gourinchas and Parker (2002). See Deaton (1992) and Attanasio (1999) for surveys.

[7] See Caballero (1990, 1991), Kimball and Mankiw (1989), Merton (1971), and Wang (2004) for explicitly-solved consumption rules with CARA utility and uninsurable labor income.

assumption of a conditionally homoskedastic Gaussian labor income process substantially simplifies the estimation of individual components of income, without losing the essential economic insight. In addition to these convenient features, this bi-variate income model is also flexible and rich enough to capture any magnitude of persistence for, and any degree of correlation between, the two individual components of income.

Unlike the optimal consumption problem when the agent has complete information about the individual components of his income (the state variables), the agent's optimization problem is no longer recursive, when he only partially observes his bi-variate income process. As a result, the original representation of the optimization problem does not land itself into a recursive structure and thus substantially complicates the analysis of the optimal consumption rule. This paper makes a methodological contribution to the consumption-saving literature by introducing a procedure to convert the original incomplete-information non-recursive optimization problem into a recursive formulation and then to derive the optimal consumption rule using dynamic programming approach. Specifically, the agent may decompose his original incomplete-information optimal consumption-saving problem into two sub-problems: (i) the signal extraction problem and (ii) an optimization problem in which he treats the estimated individual components of income as underlying state variables for his bi-variate income. I show that this two-step procedure is optimal, because the dynamics for the estimated individual components of income contains identical information as the original non-recursive incomplete information setting does. That the two-step procedure is optimal and equivalent to the original non-recursive optimization problem under incomplete information is known as the separation principle. The Kalman filter technique used in the signal extraction process allows us to derive an explicitly solved dynamics for the optimal estimate of individual components of income and thus substantially simplifies our analysis. While I have introduced the two-step methodology by working with a bi-variate income process (with total income observed), the methodology may also be applied to other settings. For example, if the agent observes his total income with noise, he then naturally also faces an optimization problem with partially observed income, the same methodology adopted in this paper may be applied to such a problem.[8]

The remainder of the paper is organized as follows. Section 2 describes the model. In Section 3, I compute the optimal consumption rule for the benchmark model in which the agent observes both components of his income. Section 4 derives the optimal consumption rule when the agent only observes total level, not individual components, of his income. Section 5 quantifies the precautionary saving premium due to estimation risk, and calculates the agent's utility loss from his ignoring the effect of estimation risk on consumption. Section 6 concludes. Appendices supply calculational details for key results of the paper.

[8]Financial economists have studied the optimal asset allocation and asset pricing under incomplete information. See Dothan and Feldman (1986), Detemple (1986), Gennotte (1986), Wang (1993), Brennan (1998), and Xia (2001), among others.

2. The model

An infinitely-lived agent receives a stream of exogenous stochastic labor income that is subject to two shocks with different degrees of persistence every period. He only observes his total income, not individual components. The agent smoothes his consumption over time by borrowing or lending at a positive constant risk-free interest rate. Thus, markets are not complete with respect to labor-income innovations. For technical convenience, I cast the model in continuous time.[9]

The agent receives labor income at time t at the rate y_t, which is equal to the sum of two components:[10]

$$y_t = w_t + z_t, \tag{1}$$

where

$$\mathrm{d}w_t = (\theta_1 - \kappa_1 w_t)\,\mathrm{d}t + \sigma_1\,\mathrm{d}B_1(t), \tag{2}$$

$$\mathrm{d}z_t = (\theta_2 - \kappa_2 z_t)\,\mathrm{d}t + \rho\sigma_2\,\mathrm{d}B_1(t) + \sqrt{1-\rho^2}\sigma_2\,\mathrm{d}B_2(t) \tag{3}$$

and ρ is the instantaneous correlation coefficient between the two individual components w and z of income. Parameters κ_1 and κ_2 measure the persistence of the respective individual component of the income process. A larger value of κ describes a less persistent process. The first-order autoregressive coefficient in a standard discrete-time formulation is equal to $e^{-\kappa}$. Positive and negative κ correspond to stationary process and non-stationary process, respectively. The unit-root process is characterized by $\kappa = 0$.

Eqs. (2) and (3) together imply that the dynamics of total income y may be written as follows:

$$\mathrm{d}y_t = (\theta - \kappa_1 w_t - \kappa_2 z_t)\,\mathrm{d}t + \mathbf{D}\,\mathrm{d}\mathbf{B}(t) = (\theta + \delta w_t - \kappa_2 y_t)\,\mathrm{d}t + \mathbf{D}\,\mathrm{d}\mathbf{B}(t), \tag{4}$$

where

$$\theta = \theta_1 + \theta_2, \tag{5}$$

$$\delta = \kappa_2 - \kappa_1, \tag{6}$$

$$v_1 = \sigma_1 + \rho\sigma_2, \tag{7}$$

$$v_2 = \sqrt{1-\rho^2}\sigma_2 \tag{8}$$

and $\mathbf{D} = (v_1\ v_2)$. If the two components have the same degrees of persistence ($\kappa_1 = \kappa_2$), then total income y is Markovian and is given by $\mathrm{d}y_t = (\theta - \kappa_1 y_t)\,\mathrm{d}t + \mathbf{D}\,\mathrm{d}\mathbf{B}(t)$. That is, there is no informational gain in estimating individual components, because the total income y is itself a univariate Markov process. The rest of this paper focuses on the more realistic and interesting case, in which the two components have

[9] A discrete-time formulation is available upon request.

[10] A standard Brownian motion $\mathbf{B} = \binom{B_1}{B_2}$ in $\mathbb{R}^2$ is given on a fixed probability space $(\Omega, \mathcal{F}, P)$, along with the standard filtration $\mathbb{F} = \{\mathcal{F}_t : t \geqslant 0\}$ of $\mathbf{B}$, where $\mathcal{F}_t$ is the information filtration generated by $\{\mathbf{B}(s) : 0 \leqslant s \leqslant t\}$, augmented by the null sets in Ω.

different degrees of persistence. Without loss of generality, I suppose that w is more persistent than z, in that $\kappa_1 < \kappa_2$. Empirical literature often models income as the sum of (i) a unit-root process and (ii) a moving-average process or a white noise process.[11] The current model offers a more flexible and richer specification for the bi-variate income process. The model not only allows for any different degrees of persistence for the individual components, but also allows for any degree of correlation between the two components of income shocks.

The agent's wealth accumulation equation is

$$\mathrm{d}x_t = (rx_t + y_t - c_t)\,\mathrm{d}t, \tag{9}$$

where $r > 0$ and x_0 is the given initial asset level. Given his information set, the agent maximizes his lifetime utility given by

$$U(c) = E\left[\int_0^\infty v(c_t,\, t)\,\mathrm{d}t\right], \tag{10}$$

subject to a transversality condition $\lim_{\tau\to\infty} E\,|\,\mathrm{e}^{-r\tau}J(x_\tau, w_\tau, z_\tau)| = 0$, where $J(x, w, z)$ is the corresponding value function for the candidate optimal consumption rule. The utility has standard exponential discounting, in that $v(c_t, t) = \mathrm{e}^{-\beta t}u(c_t)$, where $\beta > 0$ is the subjective discount rate. I choose CARA utility, in that $u(c) = -\mathrm{e}^{-\gamma c}/\gamma$ with $\gamma > 0$. The key message of the model is that estimation risk gives rise to additional precautionary saving. The CARA utility specification allows us to deliver this result in an analytically tractable way.

Following Friedman (1957) and Hall (1978), I define *human wealth* as the expected present value of future labor income discounted at the risk-free interest rate r, in that

$$h_t = E\left(\int_t^\infty \mathrm{e}^{-r(s-t)} y_s\,\mathrm{d}s \,\middle|\, \mathscr{G}_t\right), \tag{11}$$

where $\mathscr{G}_t$ is the agent's information set at time t. Note that the above definition of *human wealth* makes no adjustment to account for income risk. It simply serves as an index that converts current and expected future income to a measure of wealth. A stock measure of wealth for income helps interpreting the optimal consumption rule in later sections. In order to ensure that *human wealth* is finite, let $r + \kappa_i > 0$ for $i = 1, 2$. With positive interest rate, any stationary or unit-root process automatically satisfies $r + \kappa > 0$.

In order to quantify the precautionary saving due to partial observability of income, I first need to derive the agent's optimal consumption rule, when he observes both components of his income. The derived consumption rule under complete information serves as a starting point for us to analyze the optimal consumption rule when income is partially observed.

[11] See MaCurdy (1982), Hall and Mishkin (1982), and Pischke (1995).

3. Benchmark: complete information

If the agent observes both components of income in that his time-t information set $\mathscr{G}_t = \mathscr{F}_t$, then the three state variables for his optimization problem are wealth x, individual components w and z of current income. Appendix A shows that the optimal consumption rule c^* is affine in these three state variables, in that

$$c_t^* = r(x_t + a_1 w_t + a_2 z_t + \bar{a} - \Pi^*), \tag{12}$$

where

$$a_1 = \frac{1}{r + \kappa_1}, \tag{13}$$

$$a_2 = \frac{1}{r + \kappa_2}, \tag{14}$$

$$\bar{a} = \frac{1}{r}\left(\frac{\beta - r}{\gamma r} + \frac{\theta_1}{r + \kappa_1} + \frac{\theta_2}{r + \kappa_2}\right), \tag{15}$$

$$\Pi^* = \frac{\gamma}{2}\left[\left(\frac{\sigma_1}{r + \kappa_1}\right)^2 + 2\rho\left(\frac{\sigma_1}{r + \kappa_1}\right)\left(\frac{\sigma_2}{r + \kappa_2}\right) + \left(\frac{\sigma_2}{r + \kappa_2}\right)^2\right]. \tag{16}$$

Eq. (12) is a generalized two-factor model of Caballero (1990, 1991). In addition to the features captured in Caballero (1991), my model allows for (i) additional factor of income; (ii) flexible correlation ρ between the two components of income; (iii) flexible parameterizations of persistence κ for both components. In order to help understand the implications of the optimal consumption rule (12), it is useful to express the optimal consumption (12) in terms of the financial and human wealth (18), in that

$$c_t^* = r(x_t + h_t^* - b^*), \tag{17}$$

where

$$h_t^* = \frac{1}{r + \kappa_1}\left(w_t + \frac{\theta_1}{r}\right) + \frac{1}{r + \kappa_2}\left(z_t + \frac{\theta_2}{r}\right) \tag{18}$$

is human wealth as defined in (11) under complete information, Π^* given in (16) measures precautionary saving demand, and

$$b^* = \Pi^* - \frac{\beta - r}{\gamma r^2}. \tag{19}$$

If the agent is neither relatively patient or impatient, in that $\beta = r$, then he has no incentive to tilt his consumption towards tomorrow or today in expectation. This gives

$$c_t^* = r(x_t + h_t^* - \Pi^*). \tag{20}$$

If we further rule out precautionary saving, then agent's consumption is given by the PIH rule and is given by

$$c_t^P = r(x_t + h_t^*). \tag{21}$$

If the agent follows the consumption rule specified in (21), he saves only in anticipation of possible future declines in labor income in the sense of Campbell (1987). Correspondingly, consumption is a martingale (Hall, 1978). The difference between (20) and the PIH rule (21) measures the precautionary saving premium in units of consumption goods when income is perfectly observable and is given by

$$c_t^P - c_t^* = r\Pi^*, \tag{22}$$

where Π^* measures precautionary saving demand. Eq. (22) confirms the intuition that the precautionary saving demand is larger, for (i) a larger coefficient γ of absolute prudence, (ii) a more volatile income shock, and (iii) a more persistent (lower κ) income process, *ceteris paribus*. The intuition for (i) and (ii) is obvious. The intuition behind (iii) is as follows. While volatility parameters σ_1 and σ_2 measure income risk (in flow terms), a precautionary forward-looking agent cares the effect of income shocks not only on the current income, but also on all the future incomes. A more persistent income shock takes a longer time to die out and thus induces a stronger precautionary saving demand, *ceteris paribus*. Furthermore, a positive (negative) correlation between the two components of income increases (decreases) the total risk exposure of income risk and thus naturally induces a larger (smaller) precautionary premium.

This section presents a bi-variate stochastic income process, which allows for differing degrees of persistence and any degree of correlation between the two individual components of income, derives analytically the optimal consumption and saving rules in an intertemporal incomplete-markets setting, and discusses economic insights behind the predictions on precautionary saving. Having understood how such a complete information model works, I next introduce partial observability of income into the model. This partial observability of the income process for the agent leads to non-trivial implications on his consumption and saving decisions. This is to which we now turn.

4. Model with incomplete information

This section relaxes the assumption that both components of the income process are observable and appeals to the more realistic assumption that only total income is observed by the agent. As a result, the agent cannot precisely separate the more persistent component from the more transitory component. The inability to perfectly separate these two components leads to estimation risk, which refers to uncertainty and error associated with estimating the individual components of income from total income. The estimation risk naturally further complicates the agent's optimization problem, because he needs to choose his consumption rule, even though he does not observe the individual components of his income. Moreover, the original optimal consumption-saving problem is not recursive, because only total income is observed and total income y by itself is not recursive. (That is, the next period income cannot be specified only as a function of the current income and the shock.) In order to make the problem solvable, we need to reformulate the problem into a recursive one. Indeed, recursive methods are at the core of modern macroeconomics and economic dynamics, which are forcefully demonstrated as the unifying principle of modern macroeconomic analysis in Ljungqvist and Sargent (2004), a leading graduate textbook in macroeconomics. (The title of the book "Recursive macroeconomic theory" highlights the importance of recursive methods.)

In order to tackle this non-recursive issue of the optimization problem caused by the partial observability of income, I propose a procedure that allows us to divide the agent's optimality problem into two sub-pieces: a "signal extraction" problem and an "optimization" problem, and then to link the two pieces together to deliver the optimal consumption-saving rule via the so-called separation principle. Importantly, the "signal extraction" step allows us to derive an *equivalent* recursive representation of income dynamics, based on the history of total income y. Using this equivalent recursive representation, I convert the non-recursive optimization problem under incomplete information to an *equivalent* optimization problem with recursive structure. Then, I use the dynamic programming approach to conveniently characterize the optimal consumption rule. The separation principle verifies the equivalence between (i) the original non-recursive optimization problem when the agent only observes his total income and (ii) the newly formulated optimization problem with recursive structure, where the Markov state variables are obtained from the signal extraction problem.

First, let me turn to the recursive formulation of the signal extraction problem.

4.1. Optimal estimation of individual components of income

The objective of this section is (i) to propose a recursive formulation that describes the dynamic evolution of the optimal estimate for the unobserved individual component of income; and (ii) to show that the dynamics for the optimal estimate contains identical information as the original information specification under which the agent observes his total income in each period. I use Kalman filter technique to derive a recursive scheme for the unobserved individual component of income from a noisy measure (total income).[12] As shown later, the optimal estimate of the unobserved state variable plays a crucial role in solving the agent's optimization problem. At this stage, we may simply focus on the recursive procedure of optimally extracting signals (individual components w_t and z_t of income) out of total current income y_t, the noisy measure of individual components of income. Appendix B provides a detailed seven-step procedure for obtaining a recursive formulation for estimating the individual components of income.

[12]A classic example of the signal extraction problem in economics literature is Lucas (1973), a seminal paper on rational expectation. Muth (1960) rationalizes the "adaptive" expectations proposed in Cagan (1956) and Friedman (1957) by posing and solving an "inverse optimal estimation" problem. Muth (1960) solves this problem using classical (non-recursive) methods. See Ljungqvist and Sargent (2004) for a textbook discussion on the discrete-time formulation of the Kalman filter technique and applications to economics problems.

First, I form the state-space representation of the individual component of income by stacking (2) and (4) together. This gives

$$\mathrm{d}\begin{pmatrix} w_t \\ y_t \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \theta \end{pmatrix} + \begin{pmatrix} -\kappa_1 & 0 \\ \delta & -\kappa_2 \end{pmatrix}\begin{pmatrix} w_t \\ y_t \end{pmatrix}\right]\mathrm{d}t + \begin{pmatrix} \sigma_1 & 0 \\ v_1 & v_2 \end{pmatrix}\mathrm{d}\begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix}, \tag{23}$$

where θ, δ, v_1, and v_2 are given in (5), (6), (7), and (8), respectively. Note that while $\binom{w_t}{y_t}$ is recursive, total income y by itself is not recursive. That is, only observing current income y does not allow us to predict the next period income directly from (23), because the agent does not observe w. Thus, when income is only partially observed, we no longer have a recursive structure for the underlying bi-variate income process conditioning on the agent's information set. This substantially complicates the agent's optimal consumption-saving problem, because we are not able to write down the Bellman equation characterizing the agent's inter-temporal trade-off for a non-recursive dynamic problem. The signal extraction problem to be introduced in this section is essential in converting the non-recursive optimization problem into a recursive one.

Before delving into the details of solving the recursive signal extraction problem, I first define a few conditional expectations that will be used.Let m_t and n_t be the conditional expectations of w_t and z_t given $\mathscr{F}_t^y$, respectively, in that $m_t = E(w_t \mid \mathscr{F}_t^y)$ and $n_t = E(z_t \mid \mathscr{F}_t^y)$. Let $\Sigma(t)$ be the forecasting variance of $\binom{w_t}{z_t}$, in that

$$\Sigma(t) \equiv \begin{pmatrix} \Sigma_{11}(t) & \Sigma_{12}(t) \\ \Sigma_{21}(t) & \Sigma_{22}(t) \end{pmatrix} = E\left[\begin{pmatrix} w_t - m_t \\ z_t - n_t \end{pmatrix}\begin{pmatrix} w_t - m_t & z_t - n_t \end{pmatrix}\right].$$

Obviously, the sum of current estimates of both components must add up to the realized total income y, in that $m_t + n_t = y_t$. Otherwise, the estimation procedure cannot be optimal. The simple identity $w_t - m_t = -(z_t - n_t)$ implies that

$$\Sigma(t) = E\left[\begin{pmatrix} w_t - m_t \\ -(w_t - m_t) \end{pmatrix}\begin{pmatrix} w_t - m_t & -(w_t - m_t) \end{pmatrix}\right] = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\Sigma_{11}(t),$$

where $\Sigma_{11}(t) = E(w_t - m_t)^2$. Therefore, $\Sigma_{11}(t)$ completely characterizes the covariance matrix $\Sigma(t)$. For this conditionally homoskedastic Gaussian income process, estimation risk is fully captured by $\Sigma_{11}(t)$.

Kalman filter signal extraction technique provides a *recursive* structure to optimally estimate the unobserved underlying state variable (the individual components w and z of income) as accurately as possible by minimizing the estimation error (in the mean square sense). For a linear Gaussian system such as the bi-variate Ornstein–Uhlenbeck income process, the optimal estimate is the best *linear* estimate. Since estimation must be based on the observables, the agent thus needs to project the changes of the unobserved state variables (individual components w and z of income) onto the changes of the observed state variable (total income y). This projection procedure gives the optimal estimate m and n for the unobserved individual components w and z of income. The difference between the *realized* level of the changes in the observed total income y and the *expected* changes in total income gives rise to the error term from this projection. This error term represents *unexpected* changes of total income, dubbed as "innovations." This innovation serves as the shock (a Brownian motion in this continuous-time setting) to the optimal estimate $\binom{m}{n}$ of the individual components of income. Casting the dynamics of the optimal estimate using this innovations process gives rise to the recursive structure of the optimal estimate, the objective of the Kalman filter technique.

Next, I construct this "innovations" process from total income y and then to represent the optimal estimates m and n in terms of this newly constructed innovations process. In Appendix B, I show that the innovations process Z is a Brownian motion and the dynamics for Z is given by

$$\mathrm{d}Z(t) = \frac{1}{v}[\mathrm{d}y_t - (\theta + \delta m_t - \kappa_2 y_t)\,\mathrm{d}t], \tag{24}$$

where

$$v = \sqrt{v_1^2 + v_2^2} = \sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}} \tag{25}$$

and $\sigma_{12} = \rho\sigma_1\sigma_2$ is the instantaneous covariance between w and z, the two individual components of income. Over a small time interval Δt, the innovation $(Z(t+\Delta t) - Z(t))$ is proportional to the *unexpected* changes in total income y, which is given by the difference between $(y_{t+\Delta t} - y_t)$ and $(\theta + \delta m_t - \kappa_2 y_t)\Delta t$, where $(y_{t+\Delta t} - y_t)$ is the *actual* changes in total income and $(\theta + \delta m_t - \kappa_2 y_t)\Delta t$ is the *expected* changes in total income. Thus, (24) captures the notion of innovations, which describe "unexpected" changes or surprises given the agent's information set. The parameter v normalizes the conditional variance of this "unexpected" changes so that Z is a Brownian motion process.[13] The innovations process Z essentially plays the role as the "shock" to the optimal estimates m and n as shown below.

In Appendix B, I show that the dynamics of the optimal estimate m_t for the unobserved individual component w_t is given by

$$\mathrm{d}m_t = (\theta_1 - \kappa_1 m_t)\,\mathrm{d}t + G_1(t)\,\mathrm{d}Z(t) \tag{26}$$

with $m_0 = E(w_0)$ and where

$$G_1(t) = \frac{1}{v}(\delta\Sigma_{11}(t) + \sigma_1^2 + \sigma_{12}). \tag{27}$$

The key ideas used in deriving dynamics for the optimal estimate m_t are (i) the optimal estimate m_t may be written as a linear combination of the history of innovations ΔZ up to time t; and (ii) the residual $(w_t - m_t)$ is orthogonal to any $\mathscr{F}_t^Z$ measurable functions. (This orthogonality condition builds on the statement that m is the optimal estimate of the unobservable individual component w of income: $m_t = E(w_t|\mathscr{F}_t^y)$ and (ii) the law of the iterated expectations.) I refer readers to Appendix B for detailed derivations for (26). Using the identity $y = m + n$, (26), and the innovations process Z given in (24), we have

$$\mathrm{d}n_t = (\theta_2 - \kappa_2 n_t)\,\mathrm{d}t + G_2(t)\,\mathrm{d}Z(t) \tag{28}$$

[13] Recall that standard Brownian motion process has a standard normal distribution over a unit time interval.

with $n_0 = E(z_0)$ and where

$$G_2(t) = \frac{1}{v}(-\delta\Sigma_{11}(t) + \sigma_2^2 + \sigma_{12}). \tag{29}$$

Eqs. (26) and (28) show that the innovations process Z drives the dynamics for both m and n. Together, Eqs. (26), (28) and (24) are often dubbed the innovations representation for $\binom{m}{n}$. We note that the functional forms for the drift of the optimal estimate m and the drift of the unobserved individual components of income w are the same and are given by $f(u) = \theta_1 - \kappa_1 u$, where $u = w$ or $u = m$. In particular, the degrees of persistence for both m and w are measured by the same parameter κ_1. However, the volatility parameters for m_t and w_t are different. While the volatility parameter for w_t is equal to σ_1, the instantaneous volatility for the optimal estimate m_t, $G_1(t)$, is in general different from σ_1 and is time-varying. Specifically, the conditional volatility $G_1(t)$ depends on the forecasting variance $\Sigma_{11}(t)$.

In Appendix B, I provide detailed derivations for the evolution of the forecasting variance $\Sigma_{11}(t)$ and show that $\Sigma_{11}(t)$ solves the following Riccati equation:

$$\frac{\mathrm{d}\Sigma_{11}(t)}{\mathrm{d}t} = -2\kappa_1\Sigma_{11}(t) + [\sigma_1^2 - G_1(t)^2] \tag{30}$$

with a given initial value $\Sigma_{11}(0) = E(w_0 - m_0)^2$. Intuitively, the first term in (30) captures the forecasting variance reduction effect due to mean reversion of the unobserved individual component w. The rate of this variance reduction is bigger for a less persistent process (a higher κ_1). The other term $[\sigma_1^2 - G_1(t)^2]$ measures the difference between time-t conditional variance for w_t and time-t conditional variance for its optimal estimate m_t. The closer these two conditional variances are, the lower the estimation error at time t, and thus a smaller effect on the evolution of the conditional forecasting variance $\Sigma_{11}(t)$. Obviously, similar comparisons (for both the drift and volatility components) between the optimal estimate n and the associated underlying un-observed component z also apply.

Over time, for a stationary income process ($\kappa_1 > 0$ and $\kappa_2 > 0$), the forecasting variance $\Sigma_{11}(t)$ converges to the steady-state level $\Sigma_{11}(\infty)$. To simplify notations for the later derivations, I will focus on the steady-state variance[14] of forecasting error. Setting $\mathrm{d}\Sigma_{11}(t)/\mathrm{d}t = 0$ gives

$$\Sigma_{11}(\infty) = \frac{1}{\delta^2}\left[\sqrt{\Theta^2 + (1-\rho^2)\sigma_1^2\sigma_2^2\delta^2} - \Theta\right], \tag{31}$$

where

$$\Theta = \kappa_1\sigma_2^2 + \kappa_2\sigma_1^2 + (\kappa_1 + \kappa_2)\rho\sigma_1\sigma_2. \tag{32}$$

At the steady state ($\Sigma(t) = \Sigma(\infty)$ for all t), the joint dynamics of the optimal estimate $\binom{m_t}{n_t}$ may be written as

$$\mathrm{d}\begin{pmatrix} m_t \\ n_t \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} -\kappa_1 & 0 \\ 0 & -\kappa_2 \end{pmatrix}\begin{pmatrix} m_t \\ n_t \end{pmatrix}\right]\mathrm{d}t + \begin{pmatrix} \Lambda_1 \\ \Lambda_2 \end{pmatrix}\mathrm{d}Z(t), \tag{33}$$

[14]The time varying conditional variance is interesting in a life-cycle model. See discussions in Section 6.

where

$$\Lambda_1 = G_1(\infty) = \frac{1}{v}(\delta\Sigma_{11}(\infty) + \sigma_1^2 + \sigma_{12}),$$
$$\Lambda_2 = G_2(\infty) = \frac{1}{v}(-\delta\Sigma_{11}(\infty) + \sigma_2^2 + \sigma_{12}).$$

It is easy to show that $v = \sqrt{v_1^2 + v_2^2} = \Lambda_1 + \Lambda_2$. The parameters Λ_1 and Λ_2 are the long-run volatility parameters for the optimal estimates m and n, respectively. While the drift functions are the same for $\binom{m}{n}$ and $\binom{w}{z}$, the stationary covariance matrix of $\binom{m}{n}$ differs from that of $\binom{w}{z}$. Let Ω be the stationary covariance matrix for $\binom{w}{z}$, and $\hat{\Omega}$ be the stationary covariance matrix for $\binom{m}{n}$ under the innovations representation, respectively. Using the properties of the bi-variate Ornstein–Uhlenbeck process gives[15]

$$\Omega \equiv \lim_{t\to\infty} E\left[\begin{pmatrix} w_t - \bar{w} \\ z_t - \bar{z} \end{pmatrix}\begin{pmatrix} w_t - \bar{w} & z_t - \bar{z} \end{pmatrix}\right] = \begin{pmatrix} \frac{\sigma_1^2}{2\kappa_1} & \frac{\sigma_{12}}{\kappa_1+\kappa_2} \\ \frac{\sigma_{12}}{\kappa_1+\kappa_2} & \frac{\sigma_2^2}{2\kappa_2} \end{pmatrix}, \tag{34}$$

$$\hat{\Omega} \equiv \lim_{t\to\infty} E\left[\begin{pmatrix} m_t - \bar{w} \\ n_t - \bar{z} \end{pmatrix}\begin{pmatrix} m_t - \bar{w} & n_t - \bar{z} \end{pmatrix}\right] = \begin{pmatrix} \frac{\Lambda_1^2}{2\kappa_1} & \frac{\Lambda_1\Lambda_2}{\kappa_1+\kappa_2} \\ \frac{\Lambda_1\Lambda_2}{\kappa_1+\kappa_2} & \frac{\Lambda_2^2}{2\kappa_2} \end{pmatrix}, \tag{35}$$

where $\bar{w} = \theta_1/\kappa_1$ is the long-run mean of w and m, and $\bar{z} = \theta_2/\kappa_2$ is the long-run mean of z and n.

Using variance decomposition, I show that the forecasting covariance matrix Σ is equal to the difference between the stationary covariance matrix Ω for the underlying state vector $\binom{w_t}{z_t}$ and the stationary covariance matrix $\hat{\Omega}$ for the underlying state vector $\binom{m_t}{n_t}$, in that $\Sigma(\infty) = \Omega - \hat{\Omega}$. In order to highlight the basic intuition behind this result implied by variance decomposition, I briefly consider the analysis for $\Sigma_{11}(\infty)$. First, note that

$$E(w_t - \bar{w})^2 = E[(w_t - m_t)^2 + (m_t - \bar{m})^2] \tag{36}$$

using $E[(w_t - m_t) \mid \mathscr{F}_t^y] = 0$, $\bar{w} = \bar{m}$, and the law of iterated expectations. Using the definition for the stationary forecasting variance $\Sigma_{11}(\infty) = \lim_{t\to\infty} E(w_t - m_t)^2$ and taking the limit $t \to \infty$ on both sides of (36) give

$$\begin{aligned}\Sigma_{11}(\infty) = \lim_{t\to\infty} E(w_t - m_t)^2 &= \lim_{t\to\infty} E(w_t - \bar{w})^2 - \lim_{t\to\infty} E(m_t - \bar{m})^2, \\ &= \Omega_{11} - \hat{\Omega}_{11} = \frac{\sigma_1^2 - \Lambda_1^2}{2\kappa_1} > 0,\end{aligned} \tag{37}$$

where the first line follows from re-arranging (36) and the second line uses results on stationary variance for the Ornstein–Uhlenbeck process. Eq. (37) implies that observing total income y helps forecasting individual component w of income, in that

[15]Interchanging limits and expectation is justified in such contexts because of dominated convergence theorem.

$\Lambda_1 < \sigma_1$, and the difference between long-run stationary variance of w and that of m is the stationary variance of estimation error $w - m$, based on only observing total income y. Similar analyses give the following simple formulae for $\Sigma_{22}(\infty)$ and $\Sigma_{12}(\infty)$:

$$\Sigma_{22}(\infty) = \frac{\sigma_2^2 - \Lambda_2^2}{2\kappa_2}, \tag{38}$$

$$\Sigma_{12}(\infty) = \frac{\sigma_{12} - \Lambda_1 \Lambda_2}{\kappa_1 + \kappa_2}. \tag{39}$$

Recall that $\Sigma_{22}(\infty) = \Sigma_{11}(\infty)$ and $\Sigma_{12}(\infty) = -\Sigma_{11}(\infty)$. Negative covariance $\Sigma_{12}(\infty)$ reflects the fact that a higher estimate m for w must imply a lower estimate n for z, for a given observed total income y.

Estimating the individual components of the underlying income process is the first step towards solving the optimal consumption problem. Next, I show that the complicated optimization problem under incomplete information may be decomposed into two steps: (i) to estimate the underlying individual components of income process; and (ii) to solve the optimization problem by treating the estimated individual components of income as state variables. I show that this two-step procedure is optimal by applying the separation principle,[16] to be introduced in the next subsection.

4.2. Separation principle

Before introducing the separation principle, I briefly restate the agent's intertemporal consumption-saving problem:

$$\sup_c E\left(\int_t^\infty \mathrm{e}^{-\beta(s-t)} u(c_s)\,\mathrm{d}s \,\middle|\, \mathscr{F}_t^y \right), \tag{40}$$

subject to the wealth accumulation equation $\mathrm{d}x_t = (rx_t + y_t - c_t)\,\mathrm{d}t$ and the transversality condition specified in Appendix A. The optimization problem is also subject to what the agent knows and observes about his income process. Specifically, the agent knows the dynamic structure of his bi-variate income process, characterized by persistence parameters κ_1 and κ_2, second-moment parameters σ_1, σ_2, and ρ, and drift parameters θ_1 and θ_2 in the income process; however he does not observe w and z, the individual components of income. He only observes the history of his total income y. The complication of the optimization problem (40) to a large extent is due to the fact that total income y is not recursive with respect to $\mathscr{F}_t^y$, the agent's information set.[17]

Fortunately, the information set $\mathscr{F}_t^y$ is equivalent to $\mathscr{F}_t^Z$, the information set (filtration) generated by $\{Z(s) : 0 \leqslant s \leqslant t\}$, where dynamics of the innovations process

[16] See Fleming and Rishel (1975) and Liptser and Shiryayev (1977) for technical details.

[17] Recall that the dynamics of total income y is given by $\mathrm{d}y_t = (\theta + \delta w_t - \kappa_2 y_t)\,\mathrm{d}t + \mathbf{D}\,\mathrm{d}\mathbf{B}(t)$. Because the dynamics of total income y depends on its own value *and* also the unobservable individual component w, the total income process y itself is not recursive.

Z is given in (24). For the ease of reference, I reproduce (33), dynamics for the optimal estimate $\binom{m_t}{n_t}$, under the innovations representation:

$$\mathrm{d}\begin{pmatrix} m_t \\ n_t \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} -\kappa_1 & 0 \\ 0 & -\kappa_2 \end{pmatrix}\begin{pmatrix} m_t \\ n_t \end{pmatrix}\right]\mathrm{d}t + \begin{pmatrix} \Lambda_1 \\ \Lambda_2 \end{pmatrix}\mathrm{d}Z(t). \tag{41}$$

Unlike total income y who dynamics is not recursive under incomplete information, the optimal estimate $\binom{m}{n}$ has a recursive dynamic structure with respect to $\mathscr{F}_t^Z$. In Section 4.1 and Appendix B, using the Kalman filter technique, I show that the dynamics for the optimal estimate $\binom{m_t}{n_t}$ is linear Gaussian (following an Ornstein–Uhlenbeck process), because the underlying unobservable state dynamics for $\binom{w_t}{z_t}$ is linear with Gaussian shocks[18] and the observation equation (for total income) is linear in the underlying state variable ($y = w + z$). Moreover, the Kalman filter technique also implies that the representation (41) contains identical information as the original state-space representation (23). Luckily, the innovations representation (41) is substantially easier to work with than the state-space representation (23), due to the recursive feature of the optimal estimate $\binom{m}{n}$ in (41).

A simple version of the separation principle states that if (i) the underlying state dynamics is a linear Gaussian process in both state and control variables, (ii) the observation equation is linear in the underlying unobservable state variables, and (iii) the agent's utility function depends on the control variable,[19] then the optimal control problem with incomplete information about the state variables may be *equivalently* solved in two stages: (i) to solve the signal extraction problem for the unobservable individual components of income as in Section 4.1, and then (ii) to derive the optimal consumption rule by treating the optimally estimated individual components of income as part of the state variables. Importantly, the two-step solution procedure is optimal.

My model is a straightforward application of the separation principle stated above. Specifically, the separation principle implies that the agent's optimization problem (40) presented at the beginning of this section may be equivalently formulated as follows:

$$\sup_c E\left(\int_t^\infty \mathrm{e}^{-\beta(s-t)}u(c_s)\,\mathrm{d}s \,\middle|\, \mathscr{F}_t^Z\right), \tag{42}$$

subject to the bi-variate process (41) for the optimal estimate $\binom{m_t}{n_t}$, the wealth accumulation equation[20] $\mathrm{d}x_t = (rx_t + m_t + n_t - c_t)\,\mathrm{d}t$, and the transversality condition specified in Appendix A.

I now provide the arguments behind the separation principle as applied to this setting. An optimization problem in general specifies the following three components: (i) the objective function; (ii) the agent's information set; and (iii) the

[18]The equivalent representation for the dynamics of $\binom{w_t}{z_t}$ in discrete time is a first-order bi-variate autoregressive process.

[19]A more general statement of the separation principle allows the objective function to depend on the underlying unobservable state variable. See Fleming and Rishel (1975) for details.

[20]Plugging the identity $y = m + n$ into the wealth accumulation equation gives $\mathrm{d}x_t = (rx_t + m_t + n_t - c_t)\,\mathrm{d}t$.

constraints and laws of motions. If I show that two optimization problems are identical in all three components, then these two optimization problems are equivalent. I now check that these three components in (a) the optimal consumption problem under incomplete information and (b) the recursive optimal consumption problem using the optimal estimate as the underlying state variables, are equivalent. First, the objective functions in (40) and (42) only depend on consumption and are the same. Second, the agent's information set $\mathscr{F}_t^y$ generated by the history of his total income is equivalent to $\mathscr{F}_t^Z$, the information set generated by the history of innovations Z. Finally, the original state-space representation (23) and the innovations representation (41) are mathematically equivalent as discussed earlier and shown in detail in Appendix B. Therefore, the non-recursive optimization problem (40) is equivalent to the recursive optimization problem (42), as implied by the separation principle.

I have intentionally chosen to present a "simple" version of the separation principle and then apply it to justify the transformation of the original non-recursive optimization problem into a recursive formulation, in order to deliver the intuition of the separation principle as applied here in the simplest possible setting. However, it is worth noting that the separation principle is more general that the "simple" version, which is stated and used in this section. For example, the agent's preference in our model only depends on consumption, the control variable. However, the separation principle allows the objective function to depend on both the control variable and the unobserved state variables. For further details on the separation principle, see Fleming and Rishel (1975) and Liptser and Shiryayev (1977).

I now apply the separation principle and explicitly solve for the optimal consumption-saving rule for the equivalent recursive optimization problem (42).

4.3. Optimal consumption

I conjecture the value function in the same way as in Section 3. This optimization problem is now recursive in three state variables: wealth x, and the optimal estimates m and n for the individual components of income. Thus, we may apply the dynamic programming approach to derive the optimal consumption-saving rule. Recall that Λ_1 and Λ_2 are the volatility parameters in the innovations representation (33). Both estimated individual components of income m_t and n_t are driven by the same one-dimensional innovation Z, and thus instantaneously, m and n are perfectly correlated, as seen from (41). Appendix A verifies that the optimal consumption rule under incomplete information takes the following form

$$c_t = r(x_t + a_1 m_t + a_2 n_t + \bar{a} - \Pi), \tag{43}$$

where $a_1 = 1/(r+\kappa_1)$, $a_2 = 1/(r+\kappa_2)$,

$$\Pi = \frac{\gamma}{2}\left(\frac{\Lambda_1}{r+\kappa_1} + \frac{\Lambda_2}{r+\kappa_2}\right)^2 \tag{44}$$

and $\bar{a}$ is given in (15).

Note that the optimal consumption rule (43) under incomplete information differs from the optimal consumption rule (12) under complete information along two dimensions: (i) the underlying state variables and (ii) the functional form of the consumption rules. The underlying state variables differ in the complete-information benchmark and in the incomplete-information model because the agent's information sets are different in these two settings. More interestingly, the consumption rules also differ (even after replacing the optimal estimates m and n with their respective underlying individual components of income w and z). This difference of consumption rules essentially accounts for the additional precautionary saving induced by estimation risk. I will return to this point and discuss it more thoroughly in the next section.

Using the law of iterated expectation, the human wealth may be calculated as follows:

$$\begin{aligned} h_t &= E_t\left(\int_t^{\infty} \mathrm{e}^{-r(s-t)} y_s \, \mathrm{d}s \middle| \mathscr{F}_t^y\right) \\ &= E[E(h_t^* \mid \mathscr{F}_t) \mid \mathscr{F}_t^y] = E(h_t^* \mid \mathscr{F}_t^y) \\ &= \frac{1}{r+\kappa_1}\left(m_t + \frac{\theta_1}{r}\right) + \frac{1}{r+\kappa_2}\left(n_t + \frac{\theta_2}{r}\right), \end{aligned} \tag{45}$$

where h_t^* is the human wealth under complete information, given in (18). The optimal consumption rule (46) may also be expressed in terms of financial and human wealth (45), in that,

$$c_t = r(x_t + h_t - b), \tag{46}$$

where

$$b = \Pi - \frac{\beta - r}{\gamma r^2} \tag{47}$$

and Π is given in (44). Therefore, $r\Pi$ is the precautionary saving demand measured in units of consumption goods, when income is partially observable, following the same line of analysis in Section 3. Recall that the precautionary saving demand is $r\Pi^*$, under complete information. Intuitively, this difference between $r\Pi$ and $r\Pi^*$ is a measure for the amount of precautionary saving due to estimation risk. The next section formalizes this intuition and analyzes the effect of incomplete information about individual components of income on the agent's consumption, wealth, and lifetime utility.

5. Estimation-risk-induced precautionary saving

In order to quantify the agent's precautionary saving due to estimation risk, I naturally need a "reference" model in which the agent does not have precautionary saving motive for estimation risk, and behaves otherwise in the same way as the agent of Section 4. These considerations lead to the following choice of the "reference" model. I suppose that the behavior of the agent in the "reference" model follows the two-step procedure: (i) derive the consumption rule *assuming* that he observes both components w and z of his income y, and then (ii) replace both w and z in (17) with their best estimates m and n, respectively. The first step is the identical analysis of Section 3. The second step uses the outcome of the signal extraction problem in Section 4.1 to evaluate the policy function (17) derived in step (i). However, in the "reference" model, the agent estimates individual components of his income, only for the purpose of evaluating his already-derived consumption rule (which is optimal only under complete information). As a result, the agent in the "reference" model ignores the component of precautionary saving due to estimation risk.[21]

Because the agent has precautionary motive (convex marginal utility) in the "reference" model, he thus saves for the usual precautionary motive.[22] The difference between the incomplete-information model of Section 4.3 and the "reference" model is that the agent in the former environment takes the estimation risk into account in making his consumption decision, while the agent in the "reference" model does not take estimation risk into account in forming his consumption rule. That is, the sequence of (i) deriving the optimal consumption rule and (ii) estimating individual components of income differs in the fully optimizing model and in the "reference" model.

As a result, the consumption rule in the "reference" model may be written as

$$\hat{c}_t = E(c_t^* \mid \mathscr{G}_t) = E(c_t^* \mid \mathscr{F}_t^y) = r(x_t + a_1 m_t + a_2 n_t + \bar{a} - \Pi^*), \tag{48}$$

where c_t^* is given by (17) and $\mathscr{F}_t^y$ is the information set generated by $\{y_s : 0 \leqslant s \leqslant t\}$. Using (12) and (45), (48) may be represented as follows:

$$\hat{c}_t = r(x_t + h_t - b^*), \tag{49}$$

where h_t is human wealth for the case of partially observed income and is given in (45), and b^* is given in (19).

The difference between (46) and (49) measures the additional saving demand, that is attributable to estimation risk. Let Ξ_t denote the difference between $\hat{c}_t$ and c_t, assuming the wealth levels are the same, in that

$$\Xi_t = \hat{c}_t - c_t = \hat{c}_t(x_t, m_t, n_t) - c(x_t, m_t, n_t).$$

Precautionary saving due to estimation risk is shown below to be constant, in that

$$\begin{aligned}\Xi_t \equiv \Xi = r(\Pi - \Pi^*) &= \frac{\gamma r}{2}[(a_1\Lambda_1 + a_2\Lambda_2)^2 - (a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2\rho a_1 a_2 \sigma_1\sigma_2)] \\ &= \frac{\gamma r}{2}[a_1^2(\Lambda_1^2 - \sigma_1^2) + a_2^2(\Lambda_2^2 - \sigma_2^2) + 2a_1a_2(\Lambda_1\Lambda_2 - \sigma_{12})]\end{aligned}$$

[21] I do not use the complete-information benchmark model as the "reference" model because the underlying state variables in complete-information and incomplete-information models are different. The proposed "reference" model in this context is as close as possible to the complete-information benchmark model.

[22] Recall that (17) is the optimal consumption rule for an agent with convex marginal utility $e^{-\gamma c}$.

$$= \gamma r[-\kappa_1 a_1^2 \Sigma_{11}(\infty) - \kappa_2 a_2^2 \Sigma_{22}(\infty) - (\kappa_1 + \kappa_2) a_1 a_2 \Sigma_{12}(\infty)]$$
$$= \gamma \Sigma_{11}(\infty) \left(\frac{r\delta}{(r+\kappa_1)(r+\kappa_2)} \right)^2, \tag{50}$$

where $\Sigma_{11}(\infty)$ is given in (31). The above derivations use (37), (38), (39), and the relationship $\Sigma_{11}(\infty) = \Sigma_{22}(\infty) = -\Sigma_{12}(\infty)$. Plugging (31) into (50) gives

$$\Xi = \gamma \left(\frac{r}{(r+\kappa_1)(r+\kappa_2)} \right)^2 \left[\sqrt{\Theta^2 + (1-\rho^2)\sigma_1^2\sigma_2^2\delta^2} - \Theta \right], \tag{51}$$

where $\Theta = \kappa_1\sigma_2^2 + \kappa_2\sigma_1^2 + (\kappa_1+\kappa_2)\rho\sigma_1\sigma_2$, as given in (32). Precautionary saving Ξ due to estimation risk is higher for a larger value of $\Sigma_{11}(\infty)$, *ceteris paribus*.

Intuitively, higher volatility of individual components of income (bigger σ_1 or σ_2) leads to a larger estimation risk, in that $\mathrm{d}\Sigma_{11}(\infty)/\mathrm{d}\sigma_i > 0$, *ceteris paribus*. Therefore, precautionary saving Ξ due to estimation risk increases in σ_1 and σ_2, in that[23]

$$\frac{\mathrm{d}\Xi}{\mathrm{d}\sigma_1} = \gamma \left(\frac{r}{(r+\kappa_1)(r+\kappa_2)} \right)^2 \frac{\delta^2(\sigma_1\Lambda_2 - \rho\sigma_2\Lambda_1)\Lambda_2}{\sqrt{\Theta^2 + (1-\rho^2)\sigma_1^2\sigma_2^2\delta^2}} > 0, \tag{52}$$

$$\frac{\mathrm{d}\Xi}{\mathrm{d}\sigma_2} = \gamma \left(\frac{r}{(r+\kappa_1)(r+\kappa_2)} \right)^2 \frac{\delta^2(\sigma_2\Lambda_1 - \rho\sigma_1\Lambda_2)\Lambda_1}{\sqrt{\Theta^2 + (1-\rho^2)\sigma_1^2\sigma_2^2\delta^2}} > 0. \tag{53}$$

A larger correlation coefficient ρ between w and z implies that y is a more precise signal of w, and therefore leads to a lower estimation risk $\Sigma_{11}(\infty)$, *ceteris paribus*, in that $\mathrm{d}\Sigma_{11}(\infty)/\mathrm{d}\rho < 0$. Thus, precautionary saving Ξ due to estimation risk shall decrease in correlation coefficient ρ, as confirmed below

$$\frac{\mathrm{d}\Xi}{\mathrm{d}\rho} = -\gamma \left(\frac{r}{(r+\kappa_1)(r+\kappa_2)} \right)^2 \frac{\delta^2\sigma_1\sigma_2\Lambda_1\Lambda_2}{\sqrt{\Theta^2 + (1-\rho^2)\sigma_1^2\sigma_2^2\delta^2}} < 0. \tag{54}$$

Obviously, a larger coefficient γ of absolute prudence implies larger precautionary saving demand Ξ due to estimation risk, *ceteris paribus*.

So far, I have assumed that the wealth level x is the same in both the full optimization model and the naive reference model. Therefore, precautionary saving measured in terms of consumption goods is constant for any t. However, I have side-stepped from the effect of wealth accumulation on consumption. Consider the following setup. Two agents A and B are endowed with the same initial level of wealth x_0 at time 0. Agent A follows the optimal consumption rule (43) from time 0

[23]The inequalities of (52) and (53) use $\Lambda_1 < \sigma_1$, $\Lambda_2 < \sigma_2$, and $\Lambda_1\Lambda_2 > \sigma_{12}$.

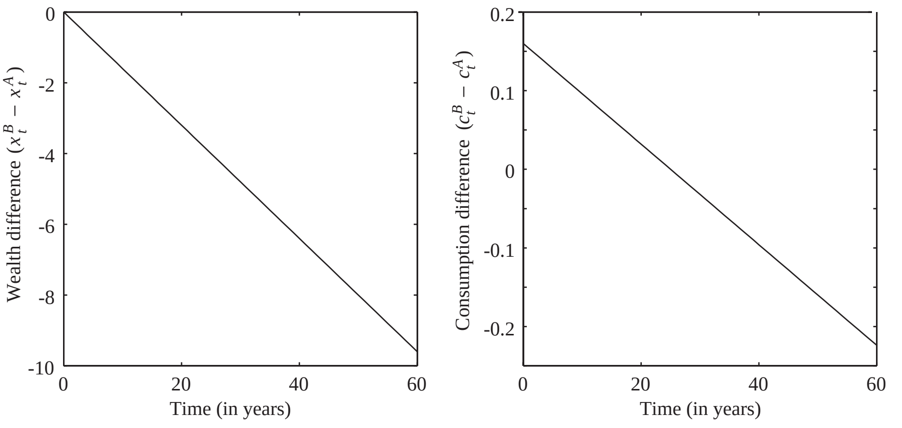

Fig. 1. Consumption and wealth differences over time.

and onward. As a result, his wealth at time t is given by

$$\begin{aligned} x_t^A &= x_0 + \int_0^t (r x_s^A + m_s + n_s - c_s)\,\mathrm{d}s \\ &= x_0 + \int_0^t [\kappa_1 a_1 m_s + \kappa_2 a_2 n_s - r(\bar{a} - \Pi)]\,\mathrm{d}s \\ &= x_0 - r(\bar{a} - \Pi)t + \int_0^t (\kappa_1 a_1 m_s + \kappa_2 a_2 n_s)\,\mathrm{d}s. \end{aligned} \tag{55}$$

Agent B lives in the "reference" model, adopting the naive consumption rule $\hat{c}$ of (48) from time 0 to t. His wealth at time t is given by

$$\begin{aligned} x_t^B &= x_0 + \int_0^t (r x_s^B + m_s + n_s - \hat{c}_s)\,\mathrm{d}s \\ &= x_0 - r(\bar{a} - \Pi^*)t + \int_0^t (\kappa_1 a_1 m_s + \kappa_2 a_2 n_s)\,\mathrm{d}s. \end{aligned} \tag{56}$$

A comparison between (55) and (56) implies

$$x_t^B - x_t^A = r(\Pi^* - \Pi)t = -\Xi t. \tag{57}$$

Over time, precautionary saving demand Ξ due to estimation risk makes Agent A, who follows the optimal consumption rule, richer and richer over time than Agent B, who does not have precautionary saving motive for estimation risk. The left diagram of Fig. 1 plots (57), the linear relationship of $(x_t^B - x_t^A)$ over time.[24]

[24] The annual interest rate is set at 4%, and the precautionary saving Ξ dues to estimation risk is chosen to be .16.

The difference between consumption of Agent A and that of Agent B at time t is given by

$$c_t^B - c_t^A = r(x_t^B - x_t^A) + \Xi = (1 - rt)\Xi. \tag{58}$$

Precautionary saving due to estimation risk has both direct and indirect effects on consumption, as seen from (58). The direct effect leads to a lower consumption level for Agent A than for Agent B, for the same wealth level ($x_t^A = x_t^B$). The indirect wealth-accumulation effect leads to a higher level of wealth for Agent A than for Agent B, as seen from (57). However, during early stages of life, there is little difference in wealth level from adopting different consumption rules. Thus, direct effect dominates the indirect effect, and $c_t^B > c_t^A$, for $t < 1/r$. After passing age $t = 1/r$, wealth of Agent A, who follows the optimal consumption rule (43), is sufficiently larger than wealth for the naive Agent B. Therefore, indirect wealth-accumulation effect has a larger impact on consumption from $t = 1/r$ and onward. The right diagram of Fig. 1 plots (58), the difference between consumption by Agent B and Agent A over time.

Fig. 1 shows the effect of precautionary saving due to estimation risk on wealth and consumption over time. Naturally, the next step is to quantify the utility loss for the agent who ignores the effect of estimation risk on consumption. Consider Agent B in "reference" model who follows the naive consumption rule (48) for all time $t \geqslant 0$. Let U^B be the lifetime utility for Agent B. Let U^A be the value function for Agent A, who takes estimation risk into account in forming his consumption decision and thus derives the optimal consumption rule (43). That is, $U^A = V(x_0, m_0, n_0)$ where $V(x, m, n)$ is given in (A.12). Appendix C shows that

$$U^B(x_0, m_0, n_0) = U^A \, \frac{1}{1 - \gamma \Xi} \, e^{-\gamma \Xi}. \tag{59}$$

Note that $e^{-\gamma\Xi} > 1 - \gamma\Xi$. Together with $U^A = V(x_0, m_0, n_0) < 0$, (59) implies $U^B < U^A$. Eq. (59) confirms the theoretical prediction that an agent following naive consumption rule has lower utility than an agent following optimal consumption rule. Next, I convert the total utility loss implied by (59) into a measure of wealth equivalent. I define the *compensating variation* η_B as the additional amount of wealth needed at time 0 in order for an agent who adopts naive consumption rule (48) to have the same utility level attainable by adopting the optimal consumption rule (43). That is, compensating variation η_B solves the following equation:

$$\begin{aligned} U^A = V(x_0, m_0, n_0) &= U^B(x_0 + \eta_B, m_0, n_0) \\ &= V(x_0 + \eta_B, m_0, n_0) \, \frac{1}{1 - \gamma \Xi} \, e^{-\gamma \Xi}, \end{aligned} \tag{60}$$

where the last equality follows from (C.13). Solving (60) gives

$$\eta_B = -\frac{1}{\gamma r}[\gamma\Xi + \log(1 - \gamma\Xi)] > 0. \tag{61}$$

Fig. 2 plots η_B against precautionary saving Ξ due to estimation risk.[25] Compensating variation η_B is increasing and convex in Ξ. Intuitively, a higher

[25]The coefficient γ of absolute risk aversion is taken to be 3, and the annual interest rate is set at 4%. The range of Ξ is from 0 to the (unattainable) upper bound $1/\gamma$, required by Condition C.12.

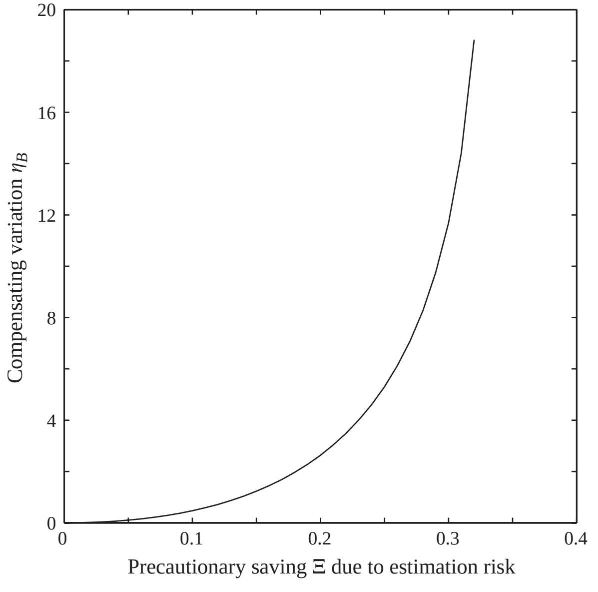

Fig. 2. Compensating variation and consumption difference over time.

precautionary saving Ξ due to estimation risk implies a bigger utility loss from ignoring the estimation risk on consumption, and therefore leads to a higher compensating variation η_B. Furthermore, a unit increase in Ξ has a larger effect on compensating variation η_B at higher level of Ξ, *ceteris paribus*.

I now compare the implications of optimal policy function (43) with those of the PIH models with partially observed income. In PIH models such as that of Pischke (1995), the agent has no precautionary saving motive at all, in that $\Pi = 0$. Therefore, estimating the individual components of income first or not yields the same consumption rule as implied by the linear-quadratic certainty-equivalence principle.[26] Suppose Agent P follows the PIH rule, and only observes his total income y. The difference between consumption of Agent A and that of Agent P is then given by

$$c^P(x_t^P, m_t, n_t) - c^A(x_t^A, m_t, n_t) = r(x_t^P - x_t^A) + \Pi = (1 - rt)\Pi, \tag{62}$$

where c^P and c^A are the PIH consumption rule and optimal consumption rule (43), respectively, and x_t^P and x_t^A are wealth levels of Agent P and Agent A at time t, respectively. By definition, Agent P and Agent A start with the same initial wealth levels x_0 at time 0. Eq. (62) states that (i) for $t \leqslant 1/r$, the optimal consumption rule (43) gives a lower consumption level than the PIH rule does, because of the agent's precautionary motive; and (ii) for $t > 1/r$, Agent A consumes more than the PIH agent, because Agent A's accumulated wealth is sufficiently large thanks to his

[26]See Ljungqvist and Sargent (2004) for a textbook treatment of certainty equivalence principle.

precautionary preference. Let U^P be the lifetime utility for Agent P, who adopts the PIH consumption rule at all times.

I define the *compensating variation* η_P as the additional amount of wealth needed at time 0 in order for an agent who adopts the PIH consumption rule (D.2) to have the same utility level attainable by adopting the optimal consumption rule (43). That is, compensating variation η_P solves the following equation:

$$U^A(x_0, m_0, n_0) = U^P(x_0 + \eta_P, m_0, n_0). \tag{63}$$

Appendix D shows that

$$U^P(x_0 + \eta_P, m_0, n_0) = U^A(x_0 + \eta_P, m_0, n_0)\,\frac{1}{1-\gamma r \Pi}\,\mathrm{e}^{-\gamma r \Pi}, \tag{64}$$

where $U^A(x_0 + \eta_P, m_0, n_0) = V(x_0 + \eta_P, m_0, n_0)$. Solving (63) and (64) jointly gives

$$\eta_P = -\frac{1}{\gamma r}[\gamma r \Pi + \log(1 - \gamma r \Pi)] > 0. \tag{65}$$

It is easy to verify the compensating variation η_P is larger than η_B.[27] The intuition is that the naive consumption rule (48) in the "reference" model only incorporates part of precautionary saving demand into account, and ignores the effect of estimation risk. As a result, adopting the naive consumption rule (48) lowers the agent's utility. The utility loss is smaller compared with the utility loss if the agent adopts the PIH rule. The intuition is straightforward. Adopting the naive consumption rule (48) ignores the effect of estimation risk on consumption, while adopting the PIH rule ignores both the effects of standard income risk and income estimation risk on consumption.

6. Conclusions

This paper studies optimal consumption and saving decisions for an agent who does not observe individual components of his income. I show that his precautionary saving demand is higher, when he optimally takes the effect of estimation risk on consumption into account. In order to show that partial observability of income leads to higher precautionary saving in a simple setting and in an intuitive way, I have chosen to use expected exponential utility and conditionally homoskedastic labor-income process. Under the current framework, the optimal approach to solve this intertemporal problem is to follow a two-step procedure: (i) estimate the underlying individual components of the income process by using Kalman filter, and (ii) derive optimal consumption rule by treating the filtered individual components of income as the underlying state variables for income process. Different from permanent-income type consumption models with unobservable income processes, such as those of Goodfriend (1992) and Pischke (1995), this paper shows that unobservability of individual components of income induces the agent to increase his precautionary saving to account for estimation risk (measured by steady-state variance $\Sigma(\infty)$ of forecasting error).

[27] Note that $H(x) = -(x + \log(1 - x))$ is an increasing function and $r\Pi > \Xi$.

This paper has focused on the infinite-horizon setting with stationary variance of estimation risk. Empirically, we observe much dispersion of individual's saving across different stages of his life cycle. One direction for future research is to incorporate the time-varying precautionary saving demand for estimation risk into a life-cycle model. Intuitively, the agent may know less about his earnings ability at the beginning of his life cycle, and learns more about his talent, a key determinant of the permanent component of his income, over time. Therefore, there is more uncertainty about the individual components of income in the earlier stage of the agent's life than in his later stage. A lower estimation risk of individual components of income in the later stage of the agent's life naturally suggests a decreasing precautionary saving demand associated with estimation risk per unit of time over his life cycle. This model may further be generalized to incorporate the effect of measurement error in the observed income on precautionary saving demand. When the observed total income is only a proxy for the underlying true income (given by the sum of the two individual components), then the estimated individual components of income is less accurate and thus gives rise to even greater demand for precautionary saving.

This paper has followed the consumption-saving literature by focusing on the effect of estimation risk on consumption-saving rules. To do so, I have intentionally chosen to allow the agent only invest in one asset, the risk-free asset with constant rate of return. We may further generalize the model to incorporate the risky asset. Because the volatility of estimated individual components of income in general differs from that of the unobserved individual components of income, estimation risk naturally also has effects on the agent's portfolio choice, which in turn has effects on the agent's consumption rule, wealth dynamics, and the utility cost of ignoring estimation risk.

Appendix A. Derivations of optimal consumption rules

First, I supply the detailed procedure of solving the optimal consumption rule when the agent has complete information about his individual components of income. Then, I solve the optimal consumption rule when only total income is observable, by applying the separation principle of Section 4.2.

A.1. Derivation of the optimal consumption rule (12)

The value function J solves the Hamilton–Jacobi–Bellman (HJB) equation:

$$0 = \sup_{\bar{c}} \{u(\bar{c}) - \beta J(x, w, z) + \mathscr{D}^{(\bar{c})} J(x, w, z)\}, \tag{A.1}$$

where

$$\begin{aligned}\mathscr{D}^{(\bar{c})} J(x, w, z) = {} & (rx + w + z - \bar{c}) J_x + (\theta_1 - \kappa_1 w) J_w + \tfrac{1}{2}\sigma_1^2 J_{ww} \\ & + (\theta_2 - \kappa_2 z) J_z + \tfrac{1}{2}\sigma_2^2 J_{zz} + \sigma_{12} J_{wz}.\end{aligned} \tag{A.2}$$

The first-order condition gives

$$e^{-\gamma\bar{c}} = J_x(x, w, z). \tag{A.3}$$

I conjecture that the value function takes an exponential-affine form

$$J(x, w, z) = -\frac{1}{\gamma r}\exp[-\gamma r(x + a_1 w + a_2 z + a_0)], \tag{A.4}$$

where a_1, a_2, and a_0 are coefficients to be determined. The FOC (A.3) implies that

$$c_t = r(x_t + a_1 w_t + a_2 z_t + a_0). \tag{A.5}$$

Plugging (A.5) into the HJB equation (A.1) gives

$$\begin{aligned} 0 = & -\frac{1}{\gamma} + \frac{\beta}{\gamma r} + [(1 - ra_1)w + (1 - ra_2)z - ra_0] + (\theta_1 - \kappa_1 w)a_1 \\ & - \frac{1}{2}\,\sigma_1^2\gamma r a_1^2 + (\theta_2 - \kappa_2 z)a_2 - \frac{1}{2}\,\sigma_2^2\gamma r a_2^2 - \gamma r a_1 a_2 \rho\sigma_1\sigma_2. \end{aligned} \tag{A.6}$$

Eq. (A.6) holds for any w and z. Setting the coefficients for w, z and constant to zero gives

$$0 = 1 - ra_1 - \kappa_1 a_1, \tag{A.7}$$

$$0 = 1 - ra_2 - \kappa_2 a_2, \tag{A.8}$$

$$0 = \frac{\beta - r}{\gamma r} - ra_0 + \theta_1 a_1 + \theta_2 a_2 - \frac{1}{2}\,\sigma_1^2\gamma r a_1^2 - \frac{1}{2}\,\sigma_2^2\gamma r a_2^2 - \gamma r a_1 a_2 \rho\sigma_1\sigma_2. \tag{A.9}$$

Solving the above three equations yield

$$a_i = \frac{1}{r + \kappa_i}, \quad i = 1, 2,$$

$$\begin{aligned} a_0 &= \frac{1}{r}\left(\frac{\beta - r}{\gamma r} + \theta_1 a_1 + \theta_2 a_2 - \frac{1}{2}\,\sigma_1^2\gamma r a_1^2 - \frac{1}{2}\,\sigma_2^2\gamma r a_2^2 - \gamma r a_1 a_2 \rho\sigma_1\sigma_2\right) \\ &= \bar{a} - \Pi^*, \end{aligned}$$

where $\bar{a}$ and Π^* are given in (15) and (16).

The transversality condition $\lim_{\tau\to\infty} E\,|\,e^{-r\tau}J(x_\tau, w_\tau, z_\tau)| = 0$, where $J(x, w, z)$ of (A.4) is the corresponding value function for the candidate optimal consumption rule, may be verified to hold in the same procedure as in Wang (2004). Details are available upon request. The key is to require the interest rate to be positive, in that $r > 0$.

A.2. Derivation of the optimal consumption rule (43)

When income is partially observable, the value function $V(x, m, n)$ solves the HJB equation:

$$0 = \sup_{\bar{c}}\{u(\bar{c}) - \beta V(x, m, n) + \mathscr{D}^{(\bar{c})}V(x, m, n)\}, \tag{A.10}$$

where

$$\mathscr{D}^{(\bar{c})}V(x,m,n) = (rx+m+n-\bar{c})V_x + (\theta_1-\kappa_1 m)V_m + \tfrac{1}{2}\Lambda_1^2 V_{mm} + (\theta_2-\kappa_2 n)V_n + \tfrac{1}{2}\Lambda_2^2\, V_{nn} + \Lambda_1\Lambda_2\, V_{mn}. \tag{A.11}$$

We follow the same procedure as in (A.1), and conjecture that the value function take the following exponential affine form

$$V(x,m,n) = -\frac{1}{\gamma r}\exp[-\gamma r(x+a_1 m+a_2 n+a_0')], \tag{A.12}$$

where a_1, a_2, and a_0' are coefficients to be determined. The implied optimal consumption rule is given by

$$c_t = r\left(x_t + \frac{1}{r+\kappa_1}\,m_t + \frac{1}{r+\kappa_2}\,n_t + a_0'\right), \tag{A.13}$$

where

$$a_0' = \frac{1}{r}\left(\frac{\beta-r}{\gamma r} + a_1\,\theta_1 + a_2\,\theta_2 - \frac{1}{2}\,\Lambda_1^2\gamma r a_1^2 - \frac{1}{2}\,\Lambda_2^2\gamma r a_2^2 - \gamma r a_1 a_2 \Lambda_1\Lambda_2\right),$$
$$= \frac{1}{r}\left[\frac{\beta-r}{\gamma r} + \frac{\theta_1}{r+\kappa_1} + \frac{\theta_2}{r+\kappa_2} - \frac{1}{2}\,\gamma r\left(\frac{\Lambda_1}{r+\kappa_1} + \frac{\Lambda_2}{r+\kappa_2}\right)^2\right] = \bar{a} - \Pi,$$

where $\bar{a}$ and Π are given in (15), and (44), respectively. Re-writing the above consumption rule gives (46), the optimal consumption rule when information is incomplete.

The same argument as for the case of perfectly observed income implies that the transversality condition holds, in that $\lim_{\tau\to\infty} E\,|\,\mathrm{e}^{-r\tau}V(x_\tau,m_\tau,n_\tau)| = 0$, where $V(x,m,n)$, given in (A.12), is the corresponding value function for the candidate optimal consumption rule.

Appendix B. Optimal estimation of individual components of income

This appendix derives the optimal estimates of the unobservable underlying individual components of income in detail using the Kalman filter technique, and shows that the innovations representation contains identical information as the original state-space dynamics does. The equivalence between the original partially observed income process and the innovations representation allows us to apply the separation principle in solving the inter-temporal optimal consumption-saving problem.

Recall that the dynamics for the individual component of income w and the associated dynamics for the observed total income y are given as follows:

$$\mathrm{d}\begin{pmatrix} w_t \\ y_t \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \theta \end{pmatrix} + \begin{pmatrix} -\kappa_1 & 0 \\ \delta & -\kappa_2 \end{pmatrix}\begin{pmatrix} w_t \\ y_t \end{pmatrix}\right]\mathrm{d}t + \begin{pmatrix} \sigma_1 & 0 \\ v_1 & v_2 \end{pmatrix}\mathrm{d}\begin{pmatrix} B_1(t) \\ B_2(t) \end{pmatrix}, \tag{B.1}$$

where $\mathbf{B}(t)$ is a standard Brownian motion in $\mathbb{R}^2$. The agent does not observe the individual component w, however, observes his total income y. His objective is to form an optimal forecast for the individual components of his income, given his information set.

I divide the detailed procedures into seven steps.

- *Step* I: Normalize shocks to the observed total income y to be one-dimensional.

In (B.1), total income y is subject to shocks from both B_1 and B_2. I normalize the shocks $\binom{B_1(t)}{B_2(t)}$ to one dimension. Let G be the Cholesky decomposition matrix for the conditional covariance matrix for $\binom{w_t}{y_t}$ from (B.1), in that

$$GG^{\mathrm{T}} = \begin{pmatrix} \sigma_1 & 0 \\ v_1 & v_2 \end{pmatrix} \begin{pmatrix} \sigma_1 & v_1 \\ 0 & v_2 \end{pmatrix}.$$

The above equation gives

$$G = \begin{pmatrix} v_2\sigma_1/v & v_1\sigma_1/v \\ 0 & v \end{pmatrix}.$$

Let $\tilde{\mathbf{B}}(t) = \begin{pmatrix} \tilde{B}_1(t) \\ \tilde{B}_2(t) \end{pmatrix}$ be another standard Brownian motion in $\mathbb{R}^2$. We may then represent the underlying state-space representation (B.1) as follows:

$$\mathrm{d}\begin{pmatrix} w_t \\ y_t \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \theta \end{pmatrix} + \begin{pmatrix} -\kappa_1 & 0 \\ \delta & -\kappa_2 \end{pmatrix} \begin{pmatrix} w_t \\ y_t \end{pmatrix} \right] \mathrm{d}t + \begin{pmatrix} v_2\sigma_1/v & v_1\sigma_1/v \\ 0 & v \end{pmatrix} \mathrm{d}\tilde{\mathbf{B}}(t). \tag{B.2}$$

Note that dynamics (B.2) for total income y depends on the past total income. This dependence complicates the signal extraction problem. Step II partially disentangles the dependence of the total income y on its own history by transforming the process $\binom{w}{y}$ to another analytically more convenient stochastic process.

- *Step* II: Simplify the state-space representation (B.2).

I first define process $\bar{w}$, whose dynamics is given by

$$\mathrm{d}\bar{w}_t = (\theta_1 - \kappa_1 \bar{w}_t)\,\mathrm{d}t + \frac{v_1\sigma_1}{v^2}[\mathrm{d}y_t - (\theta + \delta\bar{w}_t - \kappa_2 y_t)\,\mathrm{d}t], \quad \bar{w}_0 = 0. \tag{B.3}$$

Note that to fully describe the dynamics of $\bar{w}$, we only need the history of total income y. Instead of using y as the observable process, I introduce process $\tilde{y}$ as the alternative observable process, whose dynamics is given by

$$\mathrm{d}\tilde{y}_t = \mathrm{d}y_t - (\theta + \delta\bar{w}_t - \kappa_2 y_t)\,\mathrm{d}t. \tag{B.4}$$

Plugging (B.2) into (B.4) gives the following representation of $\tilde{y}$ using $\tilde{B}_2$ as the shock process:

$$\mathrm{d}\tilde{y}_t = \delta\tilde{w}_t\,\mathrm{d}t + v\,\mathrm{d}\tilde{B}_2(t).$$

Define $\tilde{w}$ to be the difference between w and $\bar{w}$, in that $\tilde{w}_t = w_t - \bar{w}_t$, as the alternative "unobservable" state process. It is straightforward to show that

$$\mathrm{d}\tilde{w}_t = -\left(\kappa_1 + \frac{v_1\sigma_1}{v^2}\,\delta\right)\tilde{w}_t\,\mathrm{d}t + \frac{v_2\sigma_1}{v}\,\mathrm{d}\tilde{B}_1(t).$$

Collecting the dynamics (B.5) for the state $\tilde{w}$ together with the observation equation (B.6) for $\tilde{y}$ gives the following new state-space representation:

$$\mathrm{d}\tilde{w}_t = -\left(\kappa_1 + \frac{v_1\sigma_1}{v^2}\,\delta\right)\tilde{w}_t\,\mathrm{d}t + \frac{v_2\sigma_1}{v}\,\mathrm{d}\tilde{B}_1(t), \tag{B.5}$$

$$\mathrm{d}\tilde{y}_t = \delta\tilde{w}_t\,\mathrm{d}t + v\,\mathrm{d}\tilde{B}_2(t). \tag{B.6}$$

Note that in (B.6), the dynamics of the observable $\tilde{y}$ does not depend on its own past, while the dynamics of the observable y depends on its own past in (B.2). From this perspective, the system (B.5)–(B.6) is simpler to work with than (B.2) does.

Let $\mathscr{F}_t^{\tilde{y}}$ be the information set (filtration) generated by the history of $\tilde{y}$, $\{\tilde{y}_s : s \leqslant t\}$, where dynamics of $\tilde{y}$ is given by (B.6). Recall that $\mathscr{F}_t^{y}$ is the information set (filtration) generated by the history of observed total income $\{y_s : s \leqslant t\}$. The next step is to show that information content in $\mathscr{F}_t^{\tilde{y}}$ is the same as the one in $\mathscr{F}_t^{y}$. Showing that $\mathscr{F}_t^{\tilde{y}}$ and $\mathscr{F}_t^{y}$ contain identical information is important because the informational equivalence between these two information sets allows us to replace the original income dynamics with a more convenient representation when solving the optimal consumption-saving rule.

- *Step* III: Show that $\mathscr{F}_t^{\tilde{y}}$ and $\mathscr{F}_t^{y}$ contain identical information.

Eq. (B.4) implies that the dynamics of $\tilde{y}$ may be obtained from the history of y and the history of $\bar{w}$.[28] Thus, we have $\mathscr{F}_t^{\tilde{y}} \subseteq \mathscr{F}_t^{y}$.

Now, we show the other direction: $\mathscr{F}_t^{y} \subseteq \mathscr{F}_t^{\tilde{y}}$. It is straightforward to express Eqs. (B.3) and (B.4) as follows:

$$\mathrm{d}\begin{pmatrix} \bar{w}_t \\ y_t \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \theta \end{pmatrix} + \begin{pmatrix} -\kappa_1 & 0 \\ \delta & -\kappa_2 \end{pmatrix}\begin{pmatrix} \bar{w}_t \\ y_t \end{pmatrix}\right]\mathrm{d}t + \begin{pmatrix} v_1\sigma_1/v^2 \\ 1 \end{pmatrix}\mathrm{d}\tilde{y}_t. \tag{B.7}$$

The first-order linear dynamics gives the following closed-form solution:

$$\begin{pmatrix} \bar{w}_t \\ y_t \end{pmatrix} = \Phi(t)\left[\begin{pmatrix} 0 \\ y_0 \end{pmatrix} + \int_0^t \Phi(s)^{-1}\begin{pmatrix} -\kappa_1 & 0 \\ \delta & -\kappa_2 \end{pmatrix}\mathrm{d}s \right.$$
$$\left. + \int_0^t \Phi(s)^{-1}\begin{pmatrix} v_1\sigma_1/v^2 \\ 1 \end{pmatrix}\mathrm{d}\tilde{y}_t\right], \tag{B.8}$$

[28]Recall that the history of $\bar{w}$ is attainable from the history of y (see (B.3)).

where[29]

$$\Phi(t) = \begin{pmatrix} e^{-\kappa_1 t} & -(1 - e^{-\kappa_1 t}) \\ e^{-\kappa_1 t} - e^{-\kappa_2 t} & e^{-\kappa_1 t} - \frac{\delta}{\kappa_2}(1 - e^{-\kappa_2 t}) \end{pmatrix}. \tag{B.9}$$

That is, information generated by $\{\tilde{y}_s : s \leqslant t\}$ allows us to solve for $\bar{w}$ and y jointly,[30] implying $\mathscr{F}_t^y \subseteq \mathscr{F}_t^{\tilde{y}}$. Thus, we have shown $\mathscr{F}_t^{\tilde{y}} = \mathscr{F}_t^y$, in that the information sets $\mathscr{F}_t^{\tilde{y}}$ and $\mathscr{F}_t^y$ contain identical information.

Next, I sketch out the basic idea behind the signal extraction problem for the "new" state-space representation given in (B.5)–(B.6): form the optimal estimate of $\tilde{w}$ given the historical observation of $\tilde{y}$.

- *Step* IV: Construct the innovations process Z for the state-space representation given in (B.5)–(B.6).

From (B.6), we may construct the following innovations representation:

$$\mathrm{d}Z(t) = \frac{1}{v}(\mathrm{d}\tilde{y}_t - \delta \tilde{m}_t \, \mathrm{d}t), \tag{B.10}$$

where I define $\tilde{m}$ as the optimal estimate of $\tilde{w}$, in that $\tilde{m}_t = E[\tilde{w}_t \,|\, \mathscr{F}_t^{\tilde{y}}]$. The basic idea behind (B.10) is simple. Start off with (B.6). Since we do not observe $\tilde{w}$, a natural thing to do is to replace $\tilde{w}$ with its optimal estimate $\tilde{m}$. Then, the deviation between $\mathrm{d}\tilde{y}$ and $\delta\tilde{m}\,\mathrm{d}t$ may then be viewed as unexpected changes, with expectation 0. In order to normalize the process Z to be a Brownian motion (which requires $\mathrm{Var}(Z(1)) = 1$), we also divide the difference $[d\tilde{y}_t - \delta\tilde{m}_t \, \mathrm{d}t]$ by the volatility parameter v. I refer readers to consult Liptser and Shiryayev (1977) for the formal proof that Z is the innovations process. Naturally, the information generated by the history of innovations process Z shall be equivalent to $\mathscr{F}_t^{\tilde{y}}$(note the relationship between Z and $\tilde{y}$ as given in (B.10)).

Next, I use the innovations process given in (B.10) to derive the innovations representation for the optimal estimate $\tilde{m}_t = E[\tilde{w}_t \,|\, \mathscr{F}_t^Z]$.

- *Step* V: Compute the innovations representation for the optimal estimate $\tilde{m}$ (for $\tilde{w}$).

The projection theorem states that we may express the optimal estimate $\tilde{m}_t$ for $\tilde{w}_t$ as a linear function of the history of innovations Z up to time t, in that

$$\tilde{m}_t = \tilde{m}_0 + \int_0^t g(s,t) \, \mathrm{d}Z(s), \tag{B.11}$$

[29]Note that the integral equation $\Phi(t)$ solves is the following:

$$\Phi(t) = I_2 + \int_0^t \begin{pmatrix} -\kappa_1 & 0 \\ \delta & -\kappa_2 \end{pmatrix} \Phi(s) \, \mathrm{d}s,$$

where I_2 is the identity matrix with 2×2.

[30]Note that $\bar{w}$ is attainable from the history of $\tilde{y}$.

where $g(s,t)$ is some well-behaved deterministic function.[31] Since $\tilde{m}$ is the "optimal" estimate of $\tilde{w}$, the residual $[\tilde{w}_t - \tilde{m}_t]$ must be orthogonal to any variable known by time t, in that

$$0 = E\left[(\tilde{w}_t - \tilde{m}_t)\int_0^t f(s)\,\mathrm{d}Z(s)\right] \tag{B.12}$$

for any well behaved function $f(\cdot)$. (Note that $\int_0^t f(s)\,\mathrm{d}Z(s)$ represents the value of a stochastic process at time t.) The orthogonality condition (B.12) implies

$$\begin{aligned} E\left[\tilde{w}_t\int_0^t f(s)\,\mathrm{d}Z(s)\right] &= E\left[\tilde{m}_t\int_0^t f(s)\,\mathrm{d}Z(s)\right] \\ &= E\left[\left(\int_0^t g(s,t)\,\mathrm{d}Z(s)\right)\left(\int_0^t f(s)\,\mathrm{d}Z(s)\right)\right] \\ &= \int_0^t f(s)g(s,t)\,\mathrm{d}s, \end{aligned} \tag{B.13}$$

where the second line follows from the application of Ito's isometry. Note that (B.13) holds for any well behaved function $f(\cdot)$. In particular, I may choose a step function: $f(s) = 0$, for $\tau \leqslant s \leqslant t$ and $f(s) = 1$, for $0 \leqslant s \leqslant \tau$. For this particular choice of $f(\cdot)$, we have $E[\tilde{w}_t Z(\tau)] = \int_0^\tau g(s,t)\,\mathrm{d}s$. Differentiating with respect to τ gives

$$g(\tau,t) = \frac{\partial}{\partial\tau}E[\tilde{w}_t Z(\tau)]. \tag{B.14}$$

Thus, the optimal estimate $\tilde{m}_t$ for $\tilde{w}_t$ is given by

$$\tilde{m}_t = E[\tilde{w}_t] + \int_0^t \frac{\partial}{\partial\tau}E[\tilde{w}_t Z(\tau)]\,\mathrm{d}Z(\tau). \tag{B.15}$$

In order to simplify (B.15), first we relate the innovations Z to the history of deviations between $\tilde{w}_s$ and its optimal estimate $\tilde{m}_s$ as follows:

$$\mathrm{d}Z(t) = \frac{1}{v}[\delta\tilde{w}_t\,\mathrm{d}t + v\,\mathrm{d}\tilde{B}_2(t) - \delta\tilde{m}_t\,\mathrm{d}t] = \frac{\delta}{v}(\tilde{w}_t - \tilde{m}_t)\,\mathrm{d}t + \mathrm{d}\tilde{B}_2(t). \tag{B.16}$$

Integrating gives

$$Z(\tau) = \frac{\delta}{v}\int_0^\tau (\tilde{w}_s - \tilde{m}_s)\,\mathrm{d}s + \tilde{B}_2(\tau). \tag{B.17}$$

Denote the variance of forecasting $\tilde{w}_t$ as $\tilde{\Sigma}_{11}(t) = E(\tilde{w}_t - \tilde{m}_t)^2$. Recall that $\tilde{w}$ is an Ornstein–Uhlenbeck process with long-run mean zero (see (B.5)), integrating (B.6) gives

$$\begin{aligned} \tilde{w}_t = {} & \exp\left[-\left(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\right)(t-s)\right]\tilde{w}_s \\ & + \frac{v_2\sigma_1}{v}\int_s^t \exp\left[-\left(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\right)(t-v)\right]\mathrm{d}\tilde{B}_1(v). \end{aligned} \tag{B.18}$$

[31] See Liptser and Shiryayev (1977) for more formal statements. The basic idea is to only consider functions in certain vector space satisfying integrability conditions.

Plugging (B.17) and (B.18) into (B.15), using the independence between $\tilde{w}$ and $\tilde{B}_2$ and orthogonality condition (B.12) give

$$E[\tilde{w}_t Z(\tau)] = \frac{\delta}{v}\int_0^\tau E[\tilde{w}_t(\tilde{w}_s - \tilde{m}_s)]\,\mathrm{d}s \tag{B.19}$$

$$= \frac{\delta}{v}\int_0^\tau E\Big[\exp\Big[-\Big(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\Big)(t-s)\Big]\tilde{w}_s(\tilde{w}_s - \tilde{m}_s)\Big]\,\mathrm{d}s, \tag{B.20}$$

$$= \frac{\delta}{v}\int_0^\tau \exp\Big[-\Big(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\Big)(t-s)\Big]\tilde{\Sigma}_{11}(s)\,\mathrm{d}s, \tag{B.21}$$

Plugging (B.21) into the integral equation (B.15) for $\tilde{m}$ gives

$$\tilde{m}_t = E[\tilde{w}_t] + \int_0^t \frac{\delta}{v}\exp\Big[-\Big(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\Big)(t-\tau)\Big]\tilde{\Sigma}_{11}(\tau)\,\mathrm{d}Z(\tau). \tag{B.22}$$

Applying Ito's formula to the above equation gives

$$\mathrm{d}\tilde{m}_t = \mathrm{d}E[\tilde{w}_t] - \Big(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\Big)(\tilde{m}_t - E[\tilde{w}_t])\,\mathrm{d}t + \frac{\delta}{v}\,\tilde{\Sigma}_{11}(t)\,\mathrm{d}Z(t) \tag{B.23}$$

$$= -\Big(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\Big)\tilde{m}_t\,\mathrm{d}t + \frac{\delta}{v}\,\tilde{\Sigma}_{11}(t)\,\mathrm{d}Z(t), \tag{B.24}$$

where the second line follows from[32]

$$\frac{\mathrm{d}}{\mathrm{d}t}\,E[\tilde{w}_t] = -\Big(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\Big)E[\tilde{w}_t]. \tag{B.25}$$

Eq. (B.24) gives the dynamics of $\tilde{m}$ up to the covariance matrix $\tilde{\Sigma}_{11}(t)$. In order to completely characterize the dynamics of the optimal estimate $\tilde{m}$ using the innovations process Z, we also need to compute the evolution of variance $\tilde{\Sigma}_{11}(t)$. This is to which we turn next.

- *Step* VI: Derive the dynamics for the forecasting variance $\tilde{\Sigma}_{11}(t) = E(\tilde{w}_t - \tilde{m}_t)^2$.

Using Ito's formula and the dynamics (B.5) for $\tilde{w}$, we have

$$\mathrm{d}\tilde{w}_t^2 = 2\tilde{w}_t\,\mathrm{d}\tilde{w}_t + \Big(\frac{v_2\sigma_1}{v}\Big)^2\mathrm{d}t \tag{B.26}$$

$$= \Big[-2\Big(\kappa_1 + \frac{v_1\sigma_1}{v^2}\,\delta\Big)\tilde{w}_t^2 + \Big(\frac{v_2\sigma_1}{v}\Big)^2\Big]\mathrm{d}t + 2\,\frac{v_2\sigma_1}{v}\,\tilde{w}_t\,\mathrm{d}\tilde{B}_1(t). \tag{B.27}$$

[32]Eq. (B.18) implies

$$E[\tilde{w}_t] = \exp\Big[-\Big(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\Big)t\Big]E[\tilde{w}_0],$$

by choosing $s = 0$ and taking expectation at time 0.

The above equation implies that the dynamics for $E[\tilde{w}_t^2]$ is given by [33]

$$\frac{\mathrm{d}}{\mathrm{d}t}E\left[\tilde{w}_t^2\right] = -2\left(\kappa_1 + \frac{v_1\sigma_1}{v^2}\delta\right)E\left[\tilde{w}_t^2\right] + \left(\frac{v_2\sigma_1}{v}\right)^2. \tag{B.28}$$

The dynamics for the second moment $\tilde{m}_t^2$ is given by

$$\mathrm{d}\tilde{m}_t^2 = 2\tilde{m}_t\,\mathrm{d}\tilde{m}_t + \left(\frac{\delta}{v}\,\tilde{\Sigma}_{11}(t)\right)^2\mathrm{d}t \tag{B.29}$$

$$= \left[-2\left(\kappa_1 + \frac{v_1\sigma_1}{v^2}\,\delta\right)\tilde{m}_t^2 + \left(\frac{\delta}{v}\,\tilde{\Sigma}_{11}(t)\right)^2\right]\mathrm{d}t + 2\,\frac{\delta}{v}\,\tilde{\Sigma}_{11}(t)\tilde{m}_t\,\mathrm{d}Z(t). \tag{B.30}$$

The above equation implies that the dynamics for $E[\tilde{m}_t^2]$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\,E[\tilde{m}_t^2] = -2\left(\kappa_1 + \frac{v_1\sigma_1}{v^2}\,\delta\right)E[\tilde{m}_t^2] + \left(\frac{\delta}{v}\,\tilde{\Sigma}_{11}(t)\right)^2. \tag{B.31}$$

Note that $\tilde{\Sigma}_{11}(t) = E(\tilde{w}_t - \tilde{m}_t)^2 = E[\tilde{w}_t]^2 - E[\tilde{m}_t]^2$. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\tilde{\Sigma}_{11}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\,E[\tilde{w}_t]^2 - \frac{\mathrm{d}}{\mathrm{d}t}[\tilde{m}_t]^2 \tag{B.32}$$

$$= -2\left(\kappa_1 + \frac{v_1\sigma_1}{v^2}\,\delta\right)\tilde{\Sigma}_{11}(t) + \left(\frac{v_2\sigma_1}{v}\right)^2 - \left(\frac{\delta}{v}\right)^2\tilde{\Sigma}_{11}(t)^2, \tag{B.33}$$

using the results from (B.28) and (B.31).

We have explicitly solved for the dynamics of the optimal estimate $\tilde{m}_t$ and its forecasting variance $\tilde{\Sigma}_{11}(t)$, using the innovations process Z. We next use results obtained from the earlier steps to derive the dynamics of the optimal estimate m_t for the unobserved individual component of income w_t and the associated forecasting variance $\Sigma_{11}(t)$.

- *Step* VII: Characterize the dynamics for the optimal estimate m_t and the dynamics for its associated forecasting variance $\Sigma_{11}(t)$.

  First, I show $\tilde{\Sigma}_{11}(t) = \Sigma_{11}(t)$. The key observation is the following identity:

$$m_t = E[w_t \mid \mathscr{F}_t^y] = E[\bar{w}_t + \tilde{w}_t \mid \mathscr{F}_t^{\tilde{y}}] = \bar{w}_t + \tilde{m}_t. \tag{B.34}$$

Therefore, we have $\tilde{w}_t - \tilde{m}_t = w_t - \bar{w}_t - \tilde{m}_t = w_t - m_t$, which implies

$$\tilde{\Sigma}_{11}(t) = E(\tilde{w}_t - \tilde{m}_t)^2 = E(w_t - m_t)^2 = \Sigma_{11}(t). \tag{B.35}$$

[33] The solution to the second moment $E[\tilde{w}_t^2]$ is given by

$$E[\tilde{w}_t^2] = \exp\left[-2\left(\kappa_1 + \frac{v_1\sigma_1}{v^2}\,\delta\right)t\right]E[\tilde{w}_0^2] + \frac{v_2^2\sigma_1^2}{2(\kappa_1 v^2 + v_1\sigma_1\delta)}\left[1 - \exp\left[-2\left(\kappa_1 + \frac{v_1\sigma_1}{v^2}\,\delta\right)t\right]\right].$$

Eq. (B.33) and $\tilde{\Sigma}_{11}(t) = \Sigma_{11}(t)$ thus give the following Riccati equation for the forecasting variance $\Sigma_{11}(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\Sigma_{11}(t) = -\frac{\delta^2}{v^2}\,\Sigma_{11}^2(t) - 2\left(\kappa_1 + \frac{\delta v_1\sigma_1}{v^2}\right)\Sigma_{11}(t) + \left(\frac{v_2\sigma_1}{v}\right)^2. \tag{B.36}$$

The solution is given by

$$\Sigma_{11}(t) = \Sigma_{11}(\infty) + \left[\frac{\delta^2}{\varepsilon v^2}(1 - \mathrm{e}^{-\varepsilon t})(\Sigma_{11}(0) - \Sigma_{11}(\infty)) + \mathrm{e}^{-\varepsilon t}\right]^{-1}(\Sigma_{11}(0) - \Sigma_{11}(\infty)),$$

where

$$\varepsilon = 2\sqrt{\kappa_1^2 + \frac{2\delta\sigma_1 v_1\kappa_1}{v^2} + \left(\frac{\delta\sigma_1}{v}\right)^2} \tag{B.37}$$

and $\Sigma_{11}(\infty)$ is the steady-state variance given in (31).

Now, I turn to the dynamics of m_t using the innovations representation. Using $m_t = \bar{w}_t + \tilde{m}_t$ from (B.34) and the dynamics for $\bar{w}$ and $\tilde{m}$ together give

$$\begin{aligned}
\mathrm{d}m_t &= (\theta_1 - \kappa_1\bar{w}_t)\,\mathrm{d}t + \frac{v_1\sigma_1}{v^2}\,\mathrm{d}\tilde{y}_t - \left(\kappa_1 + \frac{v_1\sigma_1}{v}\,\delta\right)\tilde{m}_t\,\mathrm{d}t + \frac{\delta}{v}\,\tilde{\Sigma}_{11}(t)\,\mathrm{d}Z(t)\\
&= \left(\theta_1 - \kappa_1 m_t - \frac{nu_1\sigma_1}{v}\,\delta\tilde{m}_t\right)\mathrm{d}t + \frac{v_1\sigma_1}{v}\left[\mathrm{d}Z(t) + \frac{\delta}{v}\,\tilde{m}_t\,\mathrm{d}t\right] + \frac{\delta}{v}\,\Sigma_{11}(t)\,\mathrm{d}Z(t)\\
&= (\theta_1 - \kappa_1 m_t)\,\mathrm{d}t + \left(\frac{v_1\sigma_1 + \delta\Sigma_{11}(t)}{v}\right)\mathrm{d}Z(t),
\end{aligned}$$

where Z is the Brownian motion under the innovations representation and is given by

$$\mathrm{d}Z(t) = \frac{1}{v}[\mathrm{d}\tilde{y}_t - \delta\tilde{m}_t\,\mathrm{d}t] = \frac{1}{v}[\mathrm{d}y_t - (\theta + \delta m_t - \kappa_2 y_t)\,\mathrm{d}t]. \tag{B.38}$$

To sum up, I have derived the dynamics for the optimal estimate $\binom{m_t}{n_t}$ for $\binom{w_t}{z_t}$ and the associated dynamics for the forecasting variance $\Sigma_{11}(t)$, which are given in (26)–(28) and (30), respectively. The main text uses these results to further derive the optimal consumption-saving rule and quantify the precautionary saving induced by estimation risk.

Appendix C. Calculation of lifetime utility in the "reference" model

The objective of this appendix is to calculate the lifetime utility of Agent B in the "reference" model. The (closed-loop) naive consumption rule implied by (48), the (open-loop) naive consumption rule, is given by

$$c_t^B = r\left(x_0 + \int_0^t [\kappa_1 a_1 m_s + \kappa_2 a_2 n_s - r(\bar{a} - \Pi^*)]\,\mathrm{d}s + a_1 m_t + a_2 n_t + \bar{a} - \Pi^*\right). \tag{C.1}$$

Let U^B denote lifetime utility for Agent B who follows sub-optimal consumption rule (48), in that

$$U^B = E\left[\int_0^\infty e^{-\beta t} u(c_t^B)\,dt\right]. \tag{C.2}$$

Plugging (C.1) into (C.2) gives

$$U^B = E\left[-\frac{1}{\gamma}\int_0^\infty \exp\left(-\int_0^t (\alpha_0 + \alpha_1 m_s + \alpha_2 n_s)\,ds\right) e^{-\psi_1 m_t - \psi_2 n_t}\,dt\right] e^{-\gamma r(x_0 + a_0)},$$

where

$$\psi_i = \gamma r a_i, \quad i = 1, 2, \tag{C.3}$$

$$\alpha_0 = \beta - \gamma r^2(\bar{a} - \Pi^*) = \beta - \gamma r^2 a_0, \tag{C.4}$$

$$\alpha_i = \gamma r a_i \kappa_i = \psi_i \kappa_i, \quad i = 1, 2. \tag{C.5}$$

Duffie et al. (2000) show that under technical regularity conditions, the following result holds for an "affine" process such as the one for $\binom{m_t}{n_t}$:

$$E[e^{-\int_0^\tau (\alpha_0 + \alpha_1 m_s + \alpha_2 n_s)\,ds} e^{-\psi_1 m_t - \psi_2 n_t}] = e^{\phi_0(\tau) + \phi_1(\tau) m_0 + \phi_2(\tau) n_0}, \tag{C.6}$$

where

$$\dot{\phi}_0(\tau) = -\alpha_0 + \sum_{i=1}^{2} \theta_i \phi_i(\tau) + \frac{1}{2}\begin{pmatrix}\phi_1(\tau) & \phi_2(\tau)\end{pmatrix}\begin{pmatrix}\Lambda_1^2 & \Lambda_1\Lambda_2 \\ \Lambda_1\Lambda_2 & \Lambda_2^2\end{pmatrix}\begin{pmatrix}\phi_1(\tau) \\ \phi_2(\tau)\end{pmatrix}, \tag{C.7}$$

$$\begin{pmatrix}\dot{\phi}_1(\tau) \\ \dot{\phi}_2(\tau)\end{pmatrix} = \begin{pmatrix}-\kappa_1 & 0 \\ 0 & -\kappa_2\end{pmatrix}\begin{pmatrix}\phi_1(\tau) \\ \phi_2(\tau)\end{pmatrix} - \begin{pmatrix}\alpha_1 \\ \alpha_2\end{pmatrix} \tag{C.8}$$

with boundary conditions $\phi_0(0) = 0$, $\phi_1(0) = -\psi_1$, and $\phi_2(0) = -\psi_2$. Solving the differential equation (C.8) gives

$$\phi_i(\tau) = e^{-\kappa_i \tau}\left[\phi_i(0) - \frac{\alpha_i}{\kappa_i}(e^{\kappa_i \tau} - 1)\right] = -\frac{\alpha_i}{\kappa_i} = -\psi_i, \quad i = 1, 2. \tag{C.9}$$

Using results from (C.9) to solve (C.7) gives

$$\phi_0(\tau) = \left[-\alpha_0 - (\theta_1\psi_1 + \theta_2\psi_2) + \frac{\gamma^2 r^2}{2}(a_1\Lambda_1 + a_2\Lambda_2)^2\right]\tau, \tag{C.10}$$

$$= (-r - \gamma r^2 \Pi^* + \gamma r^2 \Pi)\tau = -r(1 - \gamma\Xi)\tau, \tag{C.11}$$

where $\Xi = r(\Pi - \Pi^*)$. We assume the following convergence condition in order to ensure that the lifetime utility of Agent B who uses the naive consumption rule (48) is finite:

$$1 < \gamma\Xi. \tag{C.12}$$

Under the convergence condition (C.12), we have

$$\begin{aligned} U^B(x_0, m_0, n_0) &= -\frac{1}{\gamma}\left(\int_0^\infty e^{\phi_0(\tau)-\gamma r a_1 m_0 - \gamma r a_2 n_0}\, d\tau\right) e^{-\gamma r(x_0+a_0)} \\ &= -\frac{1}{\gamma}\, e^{-\gamma r(x_0+a_1 m_0 + a_2 n_0 + a_0)} \int_0^\infty e^{-r(1-\gamma\,\Xi)\tau}\, d\tau \\ &= -\frac{1}{\gamma}\, e^{-\gamma r(x_0+a_1 m_0 + a_2 n_0 + a_0')}\, \frac{1}{r(1-\gamma\Xi)}\, e^{-\gamma r(a_0 - a_0')} \\ &= V(x_0, m_0, n_0)\frac{1}{1-\gamma\Xi}\, e^{-\gamma\Xi} \\ &= U^A(x_0, m_0, n_0)\frac{1}{1-\gamma\Xi}\, e^{-\gamma\Xi}, \end{aligned} \tag{C.13}$$

where $V(x, m, n)$ is given in (A.12), and $a_0 - a_0' = \Pi - \Pi^*$. The above derivation uses (C.6) and the Fubini's theorem.

Appendix D. Calculation of lifetime utility in the permanent-income model

The objective of this appendix is to calculate the lifetime utility of Agent P, who forms his consumption decision based on the permanent-income hypothesis of Friedman (1957). For this bi-variate model of income, a PIH consumption rule is given by

$$c_t^P = r(x_t + a_1 m_t + a_2 n_t + \bar{a}). \tag{D.1}$$

The implied wealth dynamics is $dx_t = (\kappa_1 a_1 m_t + \kappa_2 a_2 n_t - r\bar{a})\, dt$. The implied consumption by the certainty-equivalence PIH rule in closed loop may be written as

$$c_t^P = r\left[x_0 + \int_0^t (\kappa_1 a_1 m_s + \kappa_2 a_2 n_s - r\bar{a})\, ds + a_1 m_t + a_2 n_t + \bar{a}\right]. \tag{D.2}$$

Let U^P denote lifetime utility for Agent P who follows this sub-optimal consumption rule (D.2), in that

$$U^P = E\left[\int_0^\infty e^{-\beta t} u(c_t^P)\, dt\right]. \tag{D.3}$$

Plugging (D.2) into (D.3) gives

$$U^P = E\left[-\frac{1}{\gamma}\int_0^\infty \exp\left(-\int_0^t (\alpha_0' + \alpha_1 m_s + \alpha_2 n_s)\, ds\right) e^{-\psi_1 m_t - \psi_2 n_t}\, dt\right] e^{-\gamma r(x_0+\bar{a})},$$

where $\alpha_0' = \beta - \gamma r^2 \bar{a}$, and $\alpha_i = \gamma r a_i \kappa_i = \psi_i \kappa_i$, for $i = 1, 2$.

The same argument for (C.6) gives the following result for agent following the PIH rule:

$$E[e^{-\int_0^\tau (\alpha_0' + \alpha_1 m_s + \alpha_2 n_s)\, ds}\, e^{-\psi_1 m_t - \psi_2 n_t}] = e^{\hat{\phi}_0(\tau) - \psi_1 m_0 - \psi_2 n_0}, \tag{D.4}$$

where

$$\hat{\phi}_0(\tau) = \left[-\alpha_0' - (\theta_1\psi_1 + \theta_2\psi_2) + \frac{\gamma^2 r^2}{2}(a_1\Lambda_1 + a_2\Lambda_2)^2\right]\tau, \tag{D.5}$$

$$= (-r + \gamma r^2 \Pi)\tau = -r(1 - \gamma r \Pi)\tau. \tag{D.6}$$

Assuming that the convergence condition (C.12) is met, we have

$$\begin{aligned} U^P = U^P(x_0, m_0, n_0) &= -\frac{1}{\gamma}\left(\int_0^\infty e^{\hat{\phi}_0(\tau) - \gamma r a_1 m_0 - \gamma r a_2 n_0} d\tau\right) e^{-\gamma r(x_0 + \bar{a})} \\ &= -\frac{1}{\gamma} e^{-\gamma r(x_0 + a_1 m_0 + a_2 n_0 + \bar{a})} \int_0^\infty e^{-r(1-\gamma r\Pi)\tau} d\tau \\ &= -\frac{1}{\gamma} e^{-\gamma r(x_0 + a_1 m_0 + a_2 n_0 + a_0')} \frac{1}{r(1-\gamma r\Pi)} e^{-\gamma r(\bar{a} - a_0')} \\ &= V(x_0, m_0, n_0) \frac{1}{1 - \gamma r \Pi} e^{-\gamma r \Pi} \\ &= U^A \frac{1}{1 - \gamma r \Pi} e^{-\gamma r \Pi}, \end{aligned} \tag{D.7}$$

where $U^A = V(x, m, n)$, and $\bar{a} - a_0' = \Pi$. The above derivation uses (D.4) and the Fubini's theorem. Note that the above integral is finite, because $\gamma r \Pi = \gamma(\Xi + r\Pi^*) > \gamma \Xi > 1$ using the convergence condition (C.12).

References

Abowd, J., Card, D.A., 1989. On the covariance structure of earnings and hours changes. Econometrica 57, 411–445.

Aiyagari, S.R., 1994. Uninsured idiosyncratic risk and aggregate saving. Quarterly Journal of Economics 109 (3), 659–684.

Attanasio, O.P., 1999. Consumption. In: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics, vol. 1B. North-Holland, Amsterdam, pp. 731–803.

Brennan, M.J., 1998. The role of learning in dynamic portfolio decisions. European Finance Review 1, 295–306.

Caballero, R.J., 1990. Consumption puzzles and precautionary savings. Journal of Monetary Economics 25 (1), 113–136.

Caballero, R.J., 1991. Earnings uncertainty and aggregate wealth accumulation. American Economic Review 81 (4), 859–871.

Cagan, P., 1956. The monetary dynamics of hyperinflation. In: Friedman, M. (Ed.), Studies in the Quantity Theory of Money. University of Chicago Press, Chicago, pp. 25–117.

Campbell, J.Y., 1987. Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis. Econometrica 55 (6), 1249–1273.

Carroll, C.D., 1997. Buffer-stock saving and the life cycle/permanent income hypothesis. Quarterly Journal of Economics 112 (1), 1–55.

Deaton, A., 1991. Saving and liquidity constraints. Econometrica 59 (5), 1221–1248.

Deaton, A., 1992. Understanding Consumption. Oxford University Press, Oxford.

Detemple, J., 1986. Asset pricing in a production economy with incomplete information. Journal of Finance 41, 383–391.

Dothan, M., Feldman, D., 1986. Equilibrium interest rates and multiperiod bonds in a partially observed economy. Journal of Finance 41, 369–382.
Duffie, D., Pan, J., Singleton, K.J., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68, 1343–1376.
Fleming, W.H., Rishel, R.W., 1975. Deterministic and Stochastic Optimal Control. Springer, Berlin.
Friedman, M., 1957. A Theory of the Consumption Function. Princeton University Press, Princeton.
Gennotte, G., 1986. Optimal portfolio choice under incomplete information. Journal of Finance 41, 733–749.
Goodfriend, M., 1992. Information-aggregation bias. American Economic Review 82, 508–519.
Gourinchas, P.-O., Parker, J.A., 2002. Consumption over the life cycle. Econometrica 70, 47–89.
Guvenen, F., 2004. Learning your earning: are labor income shocks really that persistent? Manuscript, University of Rochester.
Hall, R.E., 1978. Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. Journal of Political Economy 91 (6), 249–265.
Hall, R.E., Mishkin, F.S., 1982. The sensitivity of consumption to transitory income: estimates from panel data on households. Econometrica 50, 461–481.
Hansen, L.P., Sargent, T.J., 1991. Rational Expectations Econometrics. Westview Press, Boulder, Colorado.
Kimball, M.S., 1990. Precautionary saving in the small and in the large. Econometrica 58 (1), 53–73.
Kimball, M.S., Mankiw, N.G., 1989. Precautionary saving and the timing of taxes. Journal of Political Economy 97, 863–879.
Leland, H., 1968. Savings and uncertainty: the precautionary demand for savings. Quarterly Journal of Economics 82, 465–473.
Liptser, R.S., Shiryayev, A.N., 1977. Statistics of Random Processes, vol. I, II. Springer, New York.
Ljungqvist, L., Sargent, T.J., 2004. Recursive Macroeconomic Theory, second ed. MIT Press, Cambridge, MA.
Lucas Jr., R.E., 1973. Some international evidence on output-inflation tradeoffs. American Economic Review 3, 326–334.
MaCurdy, T.E., 1982. The use of time series processes to model the error structure of earnings in a longitudinal data analysis. Journal of Econometrics 18, 83–114.
Merton, R.C., 1971. Optimal consumption and portfolio rules in a continuous-time model. Journal of Economic Theory 3, 373–413.
Muth, J., 1960. Optimal properties of exponentially weighted forecasts. Journal of American Statistical Association 55, 299–306.
Pischke, J., 1995. Individual income, incomplete information, and aggregate consumption. Econometrica 63, 805–840.
Quah, D., 1990. Permanent and transitory movements in labor income: an explanation for "excess smoothness" in consumption. Journal of Political Economy 98, 449–475.
Storesletten, K., Telmer, C.I., Yaron, A., 2004. Consumption and risk sharing over the life cycle. Journal of Monetary Economics 51, 609–633.
Wang, J., 1993. A model of intertemporal asset prices under asymmetric information. Review of Economic Studies 60, 249–282.
Wang, N., 2003. Caballero meets Bewley: the permanent-income hypothesis in general equilibrium. American Economic Review 93, 927–936.
Wang, N., 2004. A simple model for Friedman's conjecture on consumption Manuscript Columbia University.
Xia, Y., 2001. Learning about predictability: the effect of parameter uncertainty on dynamic portfolio allocation. Journal of Finance 56, 205–246.
Zeldes, S.P., 1989. Optimal consumption with stochastic income: deviations from certainty equivalence. Quarterly Journal of Economics 104 (2), 275–298.