ISSN 1503-299X

WORKING PAPER SERIES

No. 5/2016

Investing it, Spending it: Interactions between Spending and Investment Decisions with a Sovereign Wealth Fund

Snorre Lindset
Department of Economics
Norwegian University of Science and Technology

Knut Anton Mork
Department of Economics
Norwegian University of Science and Technology

Department of Economics

Norwegian University of Science and Technology
N-7491 Trondheim, Norway
http://www.ntnu.edu/econ/working-papers

Investing it, Spending it: Interactions between Spending and Investment Decisions with a Sovereign Wealth Fund

November 8, 2016

Abstract

We analyze spending and investment decisions for endowments, with a special focus on Sovereign Wealth Funds. The preferred spending pattern has implications for the investment decision and the two decisions should not be separated. When interest rates deviate from their long-term mean level, this affects both the draw rate on the fund and asset allocation. The speed at which the interest rates revert back to their mean level is important. We show that short-term smoothing of the spending supported by the funds immediately increases endowment volatility because it affects the funds' principal value. Furthermore, it also tends to increase future spending volatility.

JEL classification: G11; G23

Keywords: Sovereign wealth funds, Portfolio choice, Spending

1 Introduction

Although most countries struggle to contain their public debt, Sovereign Wealth Funds (SWFs) have gained popularity in recent years. Saudi Arabia and other states in the Persian Gulf started to accumulate reserves after OPEC raised oil prices in the 1970s. The Asian financial crisis in 1998 motivated several countries to build large foreign-exchange reserves, helped by huge current-account surpluses. China has converted most of these reserves into two regular SWFs, independent of FX management. The state of Singapore manages no less than three funds to safeguard future pensions and as general financial buffers. Chile has built up a copper fund to allow the government to smooth over fluctuations in the price of copper, the country's leading export item as well as source of government revenue.

This paper is motivated by the political debate around the management of the Norwegian SWF as well as the rule for spending of the proceeds. We believe, however, that these issues have wider relevance to other SWFs as well as the endowments of non-profit institutions. We give particular attention to the interaction between how the fund is invested and how the proceeds are spent. An important result is that these two decisions cannot, in general, be separated.

The Norwegian government decided in 1990 to channel its oil and gas revenues into a special fund, initially labelled the Oil Fund, currently named the Government Pension Fund Global (GPFG, although it is in no way tied to the public pension system). As the Norwegian government has participated actively in oil and gas field investments, its net cash flow from these activities has at times been negative. The positive net revenues in the late 1970s and early 1980s were used to pay down legacy debt. With the low oil prices of the late 1980s, the government's net oil and gas revenues dropped to almost zero, so political attention was focused elsewhere. When the decision to establish the fund nevertheless was made in 1990s, the fund was not expected to become very big. With the substantial expansion of Continental Shelf activity, the development of new technology, and the high oil prices in the 21st century, it has, however, grown to its current size of NOK 7.4 trillion, corresponding to USD 845 billion at the current exchange rate. At first, investment of the fund was limited to fixed-income securities. As the fund grew in value, however, mandates for risk taking were expanded. The current mandate, given by Parliament, is based on a benchmark index with 60% equity and 40% fixed income. The mandate, furthermore, calls for regular rebalancing of the portfolio in response to security price changes.

The motivation for establishing the fund was partly to protect against Dutch decease and partly to preserve the wealth for future generations. At first, while the fund was still small, the former motivation seems to have dominated, as no plans were made for the eventual spending of the money. As the fund grew in size, however, political pressures to use it for pressing needs started to build. The response to these pressures took the form of the Norwegian Fiscal Rule of 2001, defined as follows:

1. The phasing in of oil revenues should be gradual so as to facilitate a smooth flow of government services.

2. For each fiscal year, the government may run a structural non-oil deficit up to an amount corresponding to the fund's normal real return, estimated as 4% of the fund's value at the beginning of the year.

3. Temporary deviations from the rule are allowed to the extent that discretionary fiscal policy is needed to smooth over cyclical fluctuations.

The first of these points makes it clear that smoothing is a fundamental motivation for the Norwegian fiscal rule. In the document presenting the rule, this desire for smoothing is made quite explicit. It is, of course, consistent with the literature on smoothing of taxes and public services, e.g. Barro (1979). Its operationalization in the second point is obviously inspired by the concept of permanent income (e.g. Hall (1978)).

However, the rule shares the by now well-known weakness of the permanent-income idea, namely, that it does not imply the degree of backward smoothing that most economic agents seem to prefer (Flavin (1985), Campbell and Deaton (1989), Carroll et al. (2000)). Applied to the Norwegian fiscal rule, the problem is especially that fluctuating financial returns of the fund may produce substantial variations in the fund's value over time. The fiscal rule then implicitly calls for proportional variations in the government's draw on the fund even though the size of the draw is held stable as a percentage of the fund's value. Two recent government reports (NOU 2015:9 and NOU 2016:20) recommend a gradual return to the 4% rule after temporary deviations. This is consistent with the way the rule has been practiced, for example, during and after the global financial crisis of 2008 – 2009.

The third part of the rule makes clear that the rule is not intended to stand in the way of discretionary fiscal policy as a countercyclical tool. The second part furthermore allows the automatic stabilizers to work in full. This becomes clear from the qualification that the 4% rule applies to the "structural," non-oil deficit rather than the unadjusted one.

The fact that this fund is intended to finance a smooth stream of expenditure makes it different from a pension fund, despite its name. In this paper, we argue that this difference should matter also for the way that the fund is invested. Strategies for investing the fund and for spending the proceeds should not be separated. In particular, today's practice of portfolio rebalancing after security price changes is not necessarily justified when the fund is obliged to fund a smooth stream of current expenditure. To the contrary, an adverse market development may indicate that risk taking should be reduced so as to safeguard the fund's ability to continue funding the smooth expenditure stream. Interestingly, this short-run smoothing will typically end up increasing the variability of prospective future consumption at longer horizons; however, this variability will be reduced somewhat by the proper modification of the investment strategy.

Our analysis uses key insights from the literature on the financial theory of optimal consumption and investment, dating back to Phelps (1962), Samuelson (1969), Merton (1969), and more recently Constantinides (1990) and Lax (2002). Our emphasis on the connection between investment and spending decisions is closely related to the theory on Asset Liability Management, as in Choudhry (2007). We naturally do not claim that policy can be based on theory alone. Theory is useful only if it is based on empirically valid premises. However, a mathematical framework can help discipline logical thinking and debate. In addition to the issues just mentioned, we use these tools to study some issues that we feel have not been covered, or not covered satisfactorily, in the political debate.

The first issue is the role of the subjective rate of time preference. Theory suggests that the optimal draw on a stock of capital should depend positively on this rate. In the context of the Norwegian fiscal rule, it can be thought of as an expression of policy makers' view of intergenerational fairness. If, for example, they consider current generations more deserving than future ones (perhaps because futures generations are expected to be wealthier), theory thus suggests that this could be a rational argument for higher draws on the SWF than the normal rate of return.

The second issue concerns risk. As stated, the fiscal rule makes no allowance for caution in the face of risk. Again, theory suggests that it should. That is, if the annual draw is to be based on the normal rate of return, this return should be risk adjusted.

A third issue concerns time variation in the riskless rate. The question then is whether the optimal draw should be determined by the current riskless rate or the "normal" riskless rate, somehow defined. The answer turns out to lie somewhere in between. Applied to the Norwegian fiscal rule, this means that a draw below 4% probably should be considered as long as short-term real interest rates lie very close to zero worldwide.

A government with a SWF naturally collects other revenue besides the returns on the fund. In particular, the rule's reference to the fiscal, non-oil deficit explicitly calls for draws on the fund to smooth over cyclical variations in tax revenue and/or entitlement programs, in addition to opening up for discretionary, countercyclical fiscal policy. This is analogous to an individual investor earning labor income besides the return on capital. Unfortunately, this case does not easily lend itself to closed-form solutions, so numerical simulation is required. Although we do not take up that challenge in this paper, we intend to do so as the next step in our research because we feel these issues hold considerable interest for the case of Norway.

The paper is organized as follows. Section 2 sets up the standard Merton model and looks at the implications for the role of normal real return. Section 3 analyses the implications of time variation in the risk-free rate, whereas Section 4 introduces expenditure smoothing in the form of habit formation as modelled by Constantinides (1990). We use this extension to study the implications of backward smoothing for risk taking, portfolio rebalancing, and the long-term variability of spending. Section 5 introduces an important issue that we expect to study as our next step, namely, the role of other government revenue, its typical procyclical pattern, automatic stabilizers, and discretionary, countercyclical fiscal policy. Section 6 concludes.

2 Risk-adjusted Return and Rational Myopia

We start by considering two implications of Merton's (1969) model of optimal spending and investment for an individual agent with time-additive preferences and an infinite horizon in continuous time. Like Merton, we assume power utility:

$$U = E_0 \int_0^\infty e^{-\rho t} \frac{c(t)^{1-\gamma}}{1-\gamma} dt, \tag{1}$$

where, as usual, $\rho > 0$ is the subjective rate of time preference, $\gamma > 0$ the relative rate of risk aversion, and c consumption. Because utility is time additive, the reciprocal $1/\gamma$ does double duty as the elasticity of intertemporal substitution. The limiting case of $\gamma = 1$ corresponds to logarithmic utility. For reasons to be given below, we will mostly assume $\gamma \geq 1$, so that this agent is fairly risk averse and not too willing to substitute consumption intertemporally. We believe this is consistent with the preferences of a typical government, including the one in Norway.

The agent has an initial level of wealth $W(0) = W_0$ and has the opportunity to split his or her investments between a safe asset, yielding a constant return r, and a risky asset yielding the uncertain return $z(t) \sim NIID(r + \mu, \sigma_z^2)$, where μ is the equity premium. Although we refer to the risky asset as equity, we actually think of it as a portfolio including a variety of risky assets, such as debt instruments with longer maturity or credit risk, real estate, infrastructure, and so on. For simplicity, we assume that the returns on the risky asset are serially uncorrelated. In our formal analysis, we assume away stock-price mean reversion (e.g. Fama and French (1988)) as well as uncertainty about long-term trends (Bansal and

Yaron (2004)), but offer some *ad hoc* comments on these issues when appropriate.

Let $\alpha(t)$ denote the share of wealth invested in the risky asset. Then, the expected portfolio return is $r+\alpha(t)\mu$ and its variance $\alpha(t)^2\sigma_z^2$, so that the budget constraint for utility maximization takes the form of the following diffusion process for wealth:

$$dW(t) = [(r + \alpha(t)\mu)W(t) - c(t)]dt + \alpha(t)\sigma_z W(t)dw_z(t), \tag{2}$$

where $w_z(t)$ is a Wiener process.

For easier notation in the following analysis, we define:

$$m = \frac{\mu}{\gamma\sigma_z^2}, \quad \bar{r} = r + m\mu, \quad \bar{\bar{r}} = \bar{r} - \frac{1}{2}\gamma m^2\sigma_z^2. \tag{3}$$

We also impose the following parameter constraint:

$$(1-\gamma)\bar{\bar{r}} < \rho < \gamma + (1-\gamma)\bar{\bar{r}}. \tag{4}$$

As it turns out, this constraint ensures that the transversality constraint is satisfied and that consumption always is positive, but less than total wealth.

As proved by Merton (1969), the solution to this optimization problem implies a constant value of the share α of the risky asset and the rate of consumption at a constant share η of wealth, where

$$\alpha(t) = m, \quad \eta = (1/\gamma)\rho + (1 - 1/\gamma)\bar{\bar{r}}. \tag{5}$$

This result allows us to make two observations of relevance for the Norwegian fiscal rule. For the first observation, we note that the formula for the optimal draw on the fund, η, expressed as a percentage of the fund's value, consists of two terms. In keeping with Giovannini and Weil (1989) and Campbell and Viceira (2002), we refer to these terms as a *myopic* and an *annuity* component, respectively. In contrast, the Norwegian fiscal rule only refers to the annuity component. The model suggests, however, that a rational government would combine both. If $\gamma \geq 1$, as we assume, this combination is a weighted average of the two components.

In our model, the weight $1/\gamma$ can be interpreted as either the reciprocal of the relative risk aversion or the elasticity of intertemporal substitution. The richer specification of Epstein-Zin preferences allows distinction between these two measures. The math becomes much more complex in that specification, and the optimal draw on the fund is no longer a linear combination of the myopic and the annuity component. However, Giovannini and Weil (1989) show that, in that specification, a person with a unitary elasticity of intertemporal substitution will, in our notation, set $\eta = \rho$ and thus behave completely myopically. Similarly, a person with a zero elasticity of intertemporal substitution will consume only the annuity component. Based on this insight, we interpret the factor $1/\gamma$ in (5) as the elasticity of substitution. Thus, a person that is reasonably willing to substitute consumption over time will set the draw on the fund at his or her subjective discount rate ρ without regard for financial returns.

It should be noted that there is nothing irrational or irresponsible about this type of behavior. It is simply an implication of the decision maker's preferences. If the subjective discount rate is higher than the normal financial return, wealth will tend to decline over time. So, although the draw will remain a constant share of the fund, the absolute rate of consumption will decline over time as the fund shrinks. With a high rate of time preference, this will be perfectly rational.

Applied to the Norwegian fiscal rule, a myopic spending rule could be justified as the reflection of a high subjective discount rate among policy makers. It could, for example, be motivated by a preference for favoring current generations relative to future ones, who may be expected to be more prosperous. We thus find that such favoring is rational provided the politicians also are willing to substitute consumption intertemporally. We furthermore expect that politicians that want to favor current generations also are willing to substitute consumption across generations.

We summarize this insight as

Observation 1. *The optimal draw rate on a SWF may rationally exceed the fund's annuity value if decision makers want to favor current generations and are willing to substitute* *consumption across generations.*

Our next observation concerns the annuity component of the optimal draw. The Norwegian fiscal rule stipulates this as the normal real rate of return, which in our model is $\bar{r}$. However, the optimal rule is to use the risk-adjusted version $\bar{\bar{r}}$. This gives us

Observation 2. *The annuity part of the optimal draw should not be based simply on the normal rate of return, but on a risk-adjusted version of this return.*

The form of the risk adjustment depends on the model. In our model it is simply the single-period portfolio variance times half the coefficient of relative risk aversion. Other models might define risk differently, for example, as permanent loss of capital in cases where tail risks carry more weight than in the normal distribution.

NOU 2016:20 recommends such a risk adjustment to be included in the Norwegian fiscal rule. NOU 2015:9 reports that The Norwegian Ministry of Finance has estimated the annual standard deviation for the total return on the SWF as 9.8%. Assuming $\gamma = 2$, this implies a risk adjustment of about 1 percentage point. Making this adjustment should thus reduce the Norwegian fiscal spending rule from an annual fund draw of 4% to one of 3%. With a fund of USD 800 billion, this adjustment corresponds to no less than USD 8 billion, or 3% of Norway's 2014 mainland GDP.

Mean reversion in stock returns would make the correction smaller, however. Mean reversion has been noted by Fama and French (1988) and Poterba and Summers (1988) and discussed further in Campbell and Viceira (2002). According to NOU 2015:9 (pp. 167-68), the Norwegian SWF exhibits sufficient mean reversion to reduce the standard deviation over a 15-year period to 2.5% per year. A standard deviation that small would thus reduce the above correction term by a factor greater than 10. It would still correspond to 0.2% of mainland GDP.

We have not specified our model to account for mean reversion, if any, and do not take a stand on this issue. Our main concern is that, when compared with the optimal rule in (5), the Norwegian rule for draws on the fund may on the one hand be too high in that it ignores risk correction. On the other hand, it may be too low if politicians find it fair to favor current generation over future ones. More generally, the Norwegian rule ignores the myopic term in (5). The normal rate of return should not be the only thing that matters. Risk and preferences matter as well.

3 Time-varying risk-free rates

Whereas the model in Section 2 assumes the risk-free rate to be constant, real-world risk-free rates typically vary over time. In two of his annual addresses, Norges Bank Governor Øystein Olsen (Olsen (2012) and Olsen (2015)) has argued that the secular decline in global real interest rates since the mid 1980s (as documented, e.g. by Summers (2013), King and Low (2014), Thwaites (2014), and Rachel and Smith (2015)) should indicate a lower draw rate on the SWF. This issue was also discussed in NOU 2015:9 and expanded on in NOU 2016:20. The latter report in particular advocated a lower draw rate as long as risk-free rates stay below their historical norm, even if a return to this norm can be expected eventually.

The math becomes somewhat messier when time variation in the riskless rate is considered. However, the conclusion is that the optimal draw should be lower than indicated by (5) when the riskless rate is lower than its long-term norm, but not as low as it would have been if the current rate were believed to last forever.

For this exercise, we keep the time-additive utility function (1). However, the diffusion process for wealth must be modified to reflect the time variability of the risk-free rate. Furthermore, we need to specify what happens to the equity premium in this case. We can think of two polar cases. In one of these, the equity premium remains a constant as in the above analysis. We now denote this constant $\bar{\mu}$. Alternatively, the expected equity return could be constant, which we denote φ, so that the equity premium becomes $\varphi - r(t)$. We encompass these polar cases by specifying the equity premium as

$$\mu = \mu\big(r(t)\big) = \lambda\bar{\mu} + (1-\lambda)(\varphi - r(t)), \quad 0 \leq \lambda \leq 1.$$

We can then write the diffusion process for wealth as

$$dW(t) = [(r(t) + \alpha(t)\mu)W(t) - c(t)]dt + \alpha(t)\sigma_z W(t) dw_z(t). \tag{6}$$

Furthermore, we specify the diffusion process for the risk-free rate as mean reverting (see e.g., Vasicek (1977)):

$$dr(t) = \theta[r^* - r(t)]dt + \sigma_r dw_r(t), \tag{7}$$

where $w_r(t)$ is a Wiener processes, possibly correlated with $w_z(t)$. Note that whenever $\lambda < 1$, future equity premia are stochastic. We realize that the real-world movements of risk-free rates of return likely are much more complex than the form specified in (7), see, for example, Rachel and Smith (2015). We chose our specification for the sake of mathematical simplicity and for highlighting the role of the expected speed of return to the long-term normal rate, expressed by the parameter $\theta > 0$.

This model turns out to be significantly more complex than Merton's. However, as shown in the Appendix, we can use the same method as Campbell and Viceira (2002) in their Chapter 5 to obtain an approximate solution around the long-term normal risk-free rate r^* as follows:

$$\alpha = \frac{\mu(r(t))}{\gamma\sigma_z^2} - \beta\eta'/\eta \tag{8}$$

and

$$\eta(r(t)) = \eta^* \exp\left\{\frac{(1 - 1/\gamma)[1 - (1 - \lambda)\alpha^*]}{\eta^* + \theta}(r(t) - r^*)\right\}. \tag{9}$$

Here η^* denotes the draw on the fund and α^* the equity share, respectively, that would be implied by formula (5) when the risk-free rate is r^*; and β is the theoretical regression coefficient of $\sigma_r w_r$ on $\sigma_z w_z$.

We assume $\beta \leq 0$ because a drop in the risk-free rate may imply higher stock valuations. With this assumption, and because we expect $\eta' > 0$, we see that (8) adds a positive term to the portfolio share of risky assets. Holding more of the risky asset now works as a dynamic hedge against a drop in the return on the risk-free asset. We furthermore note in the case where the equity premium is stochastic ($\lambda < 1$), that the optimal share of the risky asset is a decreasing function of the risk-free rate. However, if the equity premium remains constant, there is no new tradeoff between risk and return and thus no effect on the optimal equity share as the risk-free rate moves over time.

These considerations give us

Observation 3. *If future riskless rates are stochastic, an argument can be made for a larger equity share as a dynamic hedge. An additional argument can be made to let the equity share move in the opposite direction of the risk-free rate, but only if the equity premium is stochastic and negatively correlated with the risk-free rate.*

We are more interested in (9). The key here is the semi-elasticity of the draw rate η with respect to the time-varying interest rate, namely

$$d \ln \eta / dr(t) = (1 - 1/\gamma)[1 - (1 - \lambda)\alpha^*]/(\eta^* + \theta). \tag{10}$$

We compare it to the comparative-static semi-elasticity of the same draw rate with respect to the constant risk-free rate implied by (5):

$$d \ln \eta / dr = (1 - 1/\gamma)/\eta^*. \tag{11}$$

We thus find:

$$0 < d \ln \eta / dr(t) < d \ln \eta / dr. \tag{12}$$

We summarize this finding as

Observation 4. *When the risk-free rate temporarily falls below (rises above) its long-term normal value, the optimal draw rate should be reduced (increased), although not as much as if the drop (rise) had been permanent.*

The main difference between the two elasticities comes from the speed of adjustment of the risk-free rate back to its long-term normal. The larger this speed, the greater the difference. If the deviation from the long-term norm is truly ephemeral ($\theta \to \infty$), the draw rate should not be adjusted at all. If, however, the risk-free rate is expected to take a long time to return to normal, the adjustment should be almost as large as if the shift in the riskless rate were permanent.

4 Habit formation

As the owner of a SWF, the government will want to use it to enhance government services and/or keep a lid on taxes. As mentioned in the introduction, Barro (1979) and others have presented good arguments that both the tax system and the stream of government services ought to be smooth. NOU 2015:9 makes an explicit argument for smoothing in the context of a SWF by recommending that any deviation from the Norwegian fiscal rule should be followed by a gradual movement back to the rule.

Such gradual moves mean that the SWF budget contribution must be smooth in both a forward and a backward direction. By forward smoothness, we mean that the government plans for a smooth path going forward, given current information. This is the usual meaning of smoothing in the consumption literature. By backward smoothness, we mean that the government wants to avoid sharp changes in services and/or taxation even if the value of the fund undergoes sharp changes. We find habit formation to be a suitable technique for modeling this feature.

The consumption literature distinguishes between external and internal habits. External habits refer to people's valuation of their own consumption relative to that of others: "keeping up with the Jones'," cf. Abel (1990) or Campbell and Cochrane (1999). Internal habits refer instead to how people tend to get used to their standard of living and derive utility only from consumption over and above that standard. We believe this interpretation of the habit concept is the most relevant for the case we are studying here because our decision maker is the government deciding for the entire nation, comparable to a representative agent.

By introducing habit formation, we leave the world of time additive utility. To avoid unnecessary complexities, we now return to the assumption of a constant risk-free rate. Following Constantinides (1990), we redefine the objective function as

$$U = E_0 \int_0^\infty e^{-\rho t} \frac{[c(t) - x(t)]^{1-\gamma}}{1-\gamma} dt, \tag{13}$$

where $x(t)$ represents the habit. Like Constantinides, we specify it as

$$x(t) = e^{-at} X_0 + b \int_0^t e^{a(s-t)} c(s) ds, \quad a \geq 0, \quad b \geq 0, \quad b < r + a. \tag{14}$$

It is easily seen that this formula implies

$$dx(t) = [bc(t) - ax(t)]dt = b[c(t) - x(t)]dt + (b - a)x(t)dt. \tag{15}$$

We interpret this specification as follows. Starting from something we can think of as the subsistence level, x_0, the consumer will always, if at all possible, choose consumption of at least this level. From then on, the person will tend to get used to any given level of consumption. In the special case of $a = b$, which is adopted by Matsen (2003), the habit level will constantly increase because consumption always needs to maintain consumption a little higher than the habit level so as to avoid negatively infinite utility. If $a > b$, we see from (15) that it is possible for the habit level to decline over time. Finally, note that the factor $(b - a)$ can be thought of as an autonomous growth rate for the habit level if consumption stays just barely above the habit.

Habits matter for this person's investment strategy and for how large a part of total wealth he or she will want to consume in the first place. Based on the assumption that optimization again is done subject to the diffusion process for wealth in (2), Constantinides (1990) shows that the optimal investment and consumption strategy is characterized by the following two equations:

$$\alpha(t) = m\Big[\frac{(W(t) - x(t))/(r + a - b))}{W(t)}\Big] \tag{16}$$

and

$$c(t) = x(t) + \frac{r + a - b}{r + a}\eta[W(t) - x(t)/(r + a - b)], \tag{17}$$

where η now again is defined as in (5).

Clearly, optimal consumption (or the draw on the fund) is no longer proportional to wealth. The consumer/investor wants first to make sure that consumption at least matches the established habit level. To that, the person adds a percentage, not of overall wealth, but of the "free" part of wealth, that is wealth over and above the level $x(t)/(r + a - b)$ needed to maintain habits that grow at their autonomous rate from the current habit level. The habit-driven person is a more cautious saver because he or she cares about the ability to maintain the established level of habitual consumption.

The share of the portfolio devoted to risky assets is also no longer constant, but proportional to the ratio of free to total wealth. Put differently, the amount of wealth invested in the risky asset is a fixed proportion of the free wealth:

$$\alpha(t)W(t) = m[W(t) - x(t)/(r + a - b)]. \tag{18}$$

The reason is that the consumer/investor needs to make sure to have enough wealth to be able to finance the habit level of consumption out of the return on risk-free assets. This level is not to be gambled with.

These results are important for a government with a SWF that seeks to smooth taxes and public services. The need to preserve the desired smoothness should make the government cautious as a financial investor. As shown by Constantinides, the coefficient of relative risk aversion is no longer γ, but $\gamma/\{1 - x(t)/[(r + a - b)W(t)]\}$, which is increasing in the habit level. Similarly, the elasticity of intertemporal substitution is $(1/\gamma)[1 - x(t)/c(t)]$, which is smaller the more consumption is habit constrained. Furthermore, if the wealth shrinks because of a low return on the risky asset, the ratio $x(t)/[(r + a - b)W(t)]$ rises because the habit level is given by past consumption. Thus, bad luck in the market for risky assets should make the government more risk averse. This result is not modified by the length of the investment horizon because the habit level needs to be supported at all times, not only "in the long run." Mean reversion may make a difference; but mean reversion does not eliminate the need to support the habit level at all times.

We summarize these insights as

Observation 5. *A wish to keep taxes and public services smooth over time should make a government with a SWF want to invest the fund more conservatively than if such smoothing was not an issue. The government's apparent risk aversion should move in the opposite direction of the market for risky assets.*

Even more importantly, we note the following implication of (18):

Observation 6. *If the government wants to maintain a smooth flow of taxes and government* *services, the rules for SWF portfolio rebalancing after asset price changes should be modified so as to safeguard the funds needed to secure this smoothness.*

Recall that, without habit formation, the risky share of the portfolio should always be the constant m. So, whenever the prices of risky assets fall, the fund should buy so much more of them so as to maintain the constant value share. Under habit formation, on the other hand, a drop in the prices of risky assets means that a larger share of the portfolio needs to be set aside to ensure that the habit level of consumption can be preserved. So, the government may in some cases need to sell part of its risky portfolio rather than buying additional quantities of the risky asset, to make sure that it can preserve the habit level of services and taxes.

As shown by Sundaresan (1989) and Constantinides (1990), habit formation reduces the short-run volatility of consumption when volatility is defined as the variance of the instantaneous log change in consumption. While perhaps not obvious, this result is clearly not surprising. Indeed, habit formation has been invoked as a mechanism to help explain the empirical smoothness of consumption.

However, this smoothness carries a price. For example, after an adverse market movement, the government may need to attack the fund's principal in order to make sure that current and future draws can at least match the habit level. Thus, although consumption is smoothed, the fund's value becomes more volatile. Not only that, the fund's variance on various horizons may rise faster with the horizon than without habit formation.

This steeper rise in the long-horizon volatility of the fund value may then be translated into a higher long-horizon volatility of consumption as well. Thus, although the short-horizon variance of consumption is lower with habit formation, the long-horizon variance will be higher for consumption as well as for the fund itself.

These insights can be gleaned from inspection of the formulae involved. However, numerical simulations presented in Figure 1 show that the variance of future log-wealth for the habit-formation case rises much more quickly with the length of the horizon than in the case without habit formation (in the figure labeled “Merton”).[1] The figure shows four different sets of parameter configurations. For all four configurations, the parameters are set so that the investor with habit preferences has the same initial portfolio composition as the investor equipped with standard power utility and a coefficient of relative risk aversion $\gamma = 2$.

The long-horizon uncertainty of consumption is illustrated in Figure 2. Here, we plot the variances of log-consumption for different time horizons and for different parameter configurations. In three of the cases, we see lower consumption volatility for “short” horizons (say, less than 20 to 50 years) for the habit case than for the Merton case. However, for longer horizons the habit investor can face far more variation in consumption. This increase comes as a consequence of the riskier wealth illustrated in Figure 1.

This exercise teaches an important lesson: Smoothness carries a price. We can smooth current consumption by using the fund as a buffer. But then we tamper with the fund's principal. In so doing, we indirectly affect future consumption and hence future habits, which in turn influence consumption even further out. Short-term convenience carries long-term costs.

Provided the habits really are part of preferences, the tradeoff between short-term smoothness and long-term uncertainty is done optimally. Figure 3 shows what would happen to the variance of log-consumption at various time horizons if the investment and spending decisions were separated so that the risky share of the portfolio were kept constant even though spending is based on the above implications of the habit model. The separation of spending and investment decisions leads to higher variability of consumption. At least as interesting is the fact that the separated rules eventually become inconsistent in all four examples, illustrated by the fact that the graph for the variance of log-consumption ends prematurely for the fixed case. This happens because keeping the risky share fixed fails to safeguard the funding of the minimal, habit-determined rates of future spending for some states of the world. Thus, the investor/consumer ends up in what Lax (2002) refers to as the insolvency range.[2]

[1] Calculations are performed using Ox, see Doornik (1999).

[2] Out of the 10,000 simulated consumption paths used to construct each of the four graphs for the variance

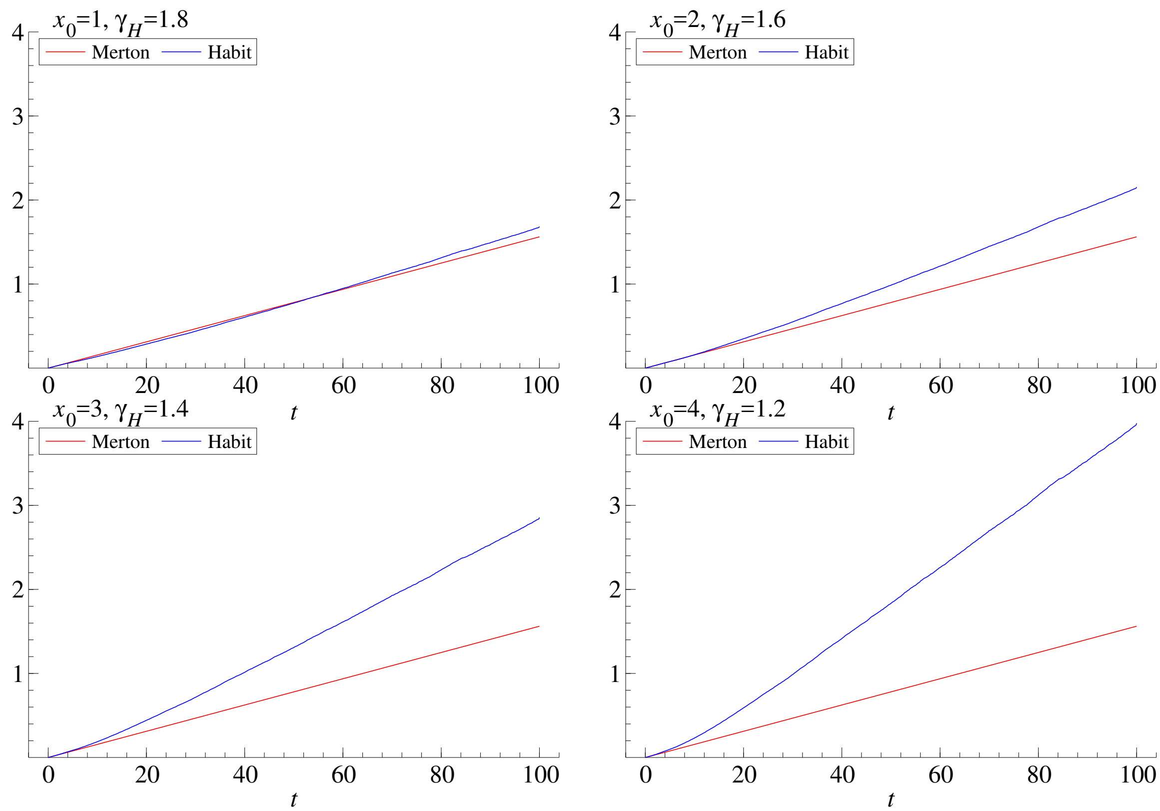

Figure 1: Variance of log-wealth at different time horizons t for the investor with power utility (Merton) and for the investor with habit preferences (Habit). Parameter values are $\mu = 0.05$, $\sigma_z = 0.20$, $r = 0.05$, $\gamma = 2$, $a = 0.3$, $b = 0.25$, $\rho = 0.03$, and $W_0 = 100$. The variances are estimated from 10,000 simulated observations at each point in time. We use 10 time points per year. The parameter x_0 shows the initial habit level and γ_H shows the γ-coefficient for the investor with habit preferences.

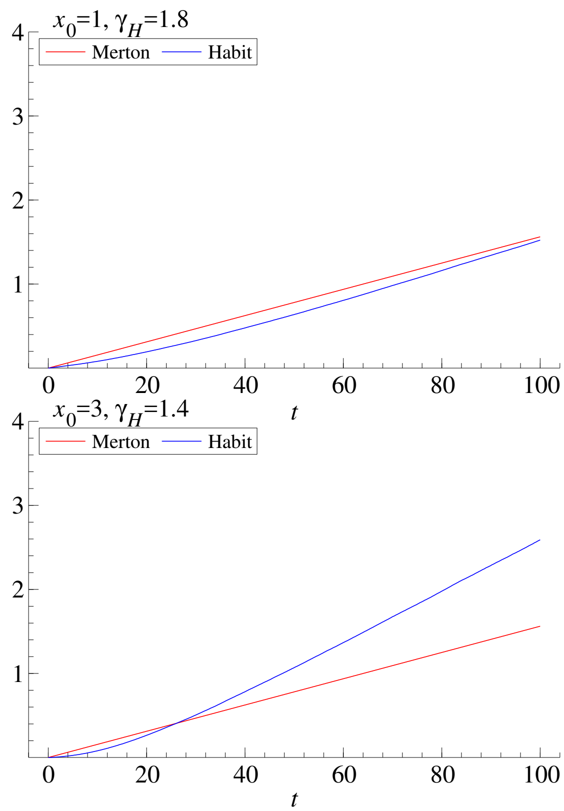

Figure 2: Variance of log-consumption at different time horizons t for the investor with power utility (Merton) and for the investor with habit preferences (Habit). Parameter values are $\mu = 0.05$, $\sigma_z = 0.20$, $r = 0.05$, $\gamma = 2$, $a = 0.3$, $b = 0.25$, $\rho = 0.03$, and $W_0 = 100$. The variances are estimated from 10,000 simulated observations at each point in time. We use 10 time points per year. The parameter x_0 shows the initial habit level and γ_H shows the γ-coefficient for the investor with habit preferences.

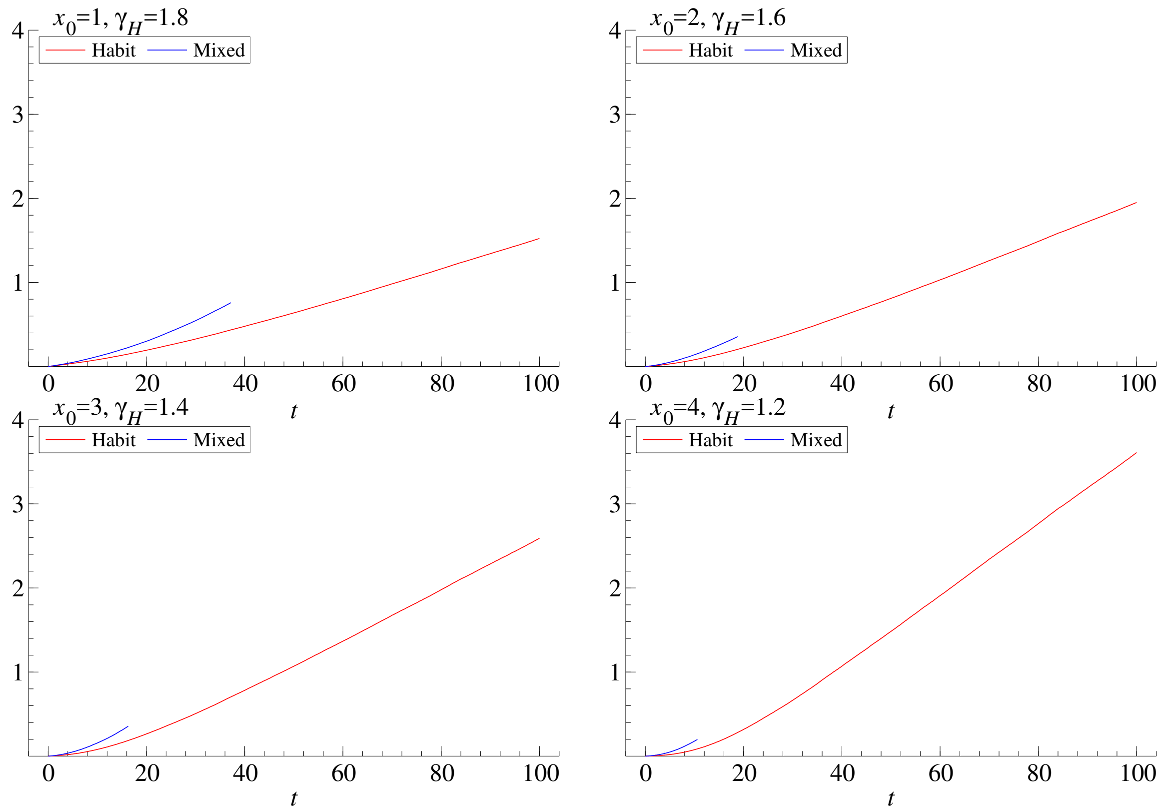

Figure 3: Variance of log-consumption at different time horizons t for the investor with habit preferences (Habit) and for a consumer following the same rules for spending, but with risky assets kept at the same portfolio share as the initial share for the Habit investor. Parameter values are $\mu = 0.05$, $\sigma_z = 0.20$, $r = 0.05$, $\gamma = 2$, $a = 0.3$, $b = 0.25$, $\rho = 0.03$, and $W_0 = 100$. The variances are estimated from 10,000 simulated observations at each point in time. We use 10 time points per year. The parameter x_0 shows the initial habit level and γ_H shows the γ-coefficient for the investor with habit preferences.

Figure 4 shows the optimal variation in the risky share. As can be seen, this variation can be quite substantial, especially at the longer horizons; and the high variance seems mainly to persist once it has become large. This persistence is consistent with the distribution of the risky-asset ratio approaching a steady state over time. Figure 5 confirms this impression by showing that the mean values tend to level out with the horizon as well. We summarize these observations as

Observation 7. *The smoothing of consumption tends to carry a price in the form of a wider uncertainty of the long-run prospects for consumption. This uncertainty is mitigated by the optimal modification of portfolio rebalancing, but it is not removed.*

5 Extensions and future work

Individuals normally receive labor income in addition to the return on their financial wealth. Likewise, most governments (with Saudi Arabia as the famous counterexample) collect tax revenue on top of the return on their financial assets even if they have a SWF. Governments usually aim to make these revenues as stable as possible over time by keeping tax rates stable. However, actual revenue naturally varies with the business cycle. Similarly, government spending, especially on entitlement programs, tends to move procyclically. Thus, in the absence of discretionary fiscal action, and ignoring the contributions from the SWF, the budget balance will tend to move countercyclically.

This, of course, is nothing but automatic stabilizers at work. They are stronger the more comprehensive the welfare state is. European governments regularly rely on this mechanism to dampen the effects of business cycles on the overall economy. Recent research by Di Maggio and Kermani (2015) indicates that it may be significant even for the United States.

However, the automatic stabilizers must be financed. The Norwegian fiscal rule implicitly calls for them to be financed from the SWF by stipulating the 4% rule in reference to the

of log-consumption (labeled "Mixed"), insolvency happens in 194, 335, 609, and 995 consumption paths, with the number increasing in x_0.

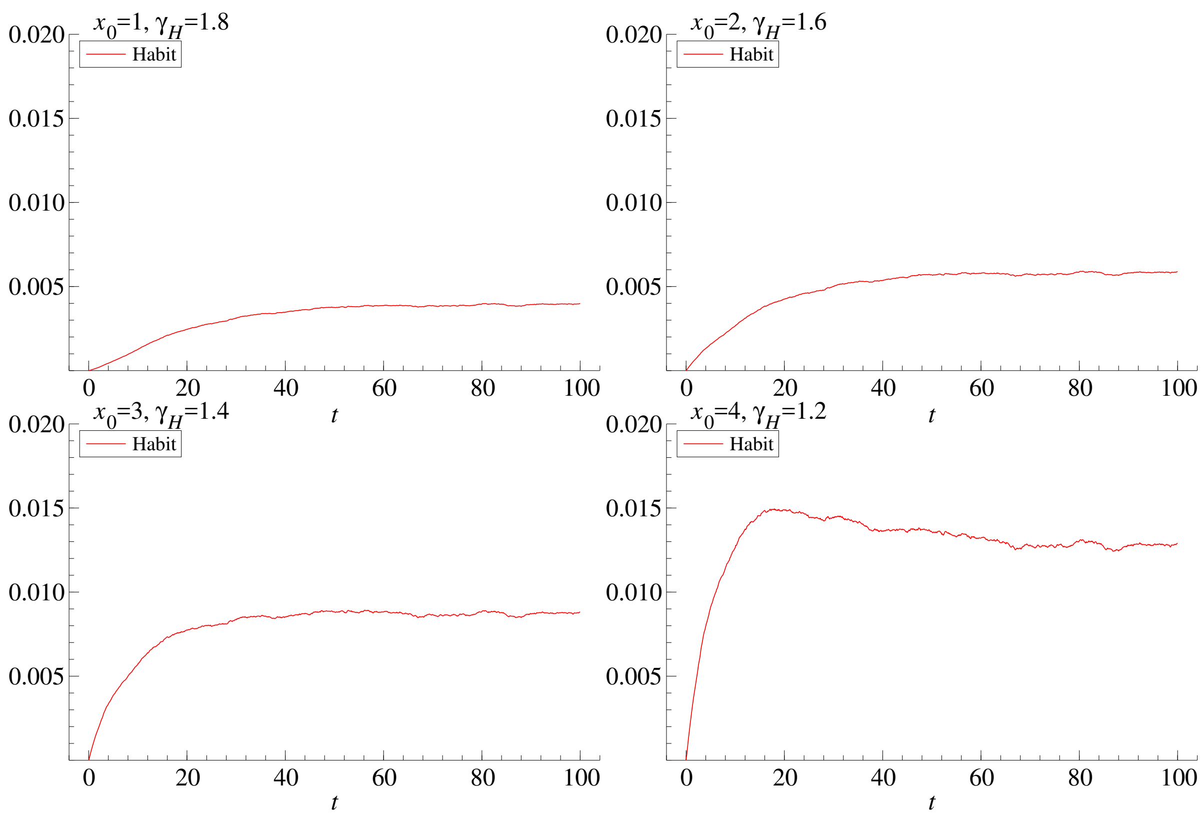

Figure 4: Optimal variance of the risky portfolio share $\alpha(t)$ under habit formation. Parameter values are $\mu = 0.05$, $\sigma_z = 0.20$, $r = 0.05$, $\gamma = 2$, $a = 0.3$, $b = 0.25$, $\rho = 0.03$, and $W_0 = 100$. The variances are estimated from 10,000 simulated observations at each point in time. We use 10 time points per year. The parameter x_0 shows the initial habit level and γ_H shows the γ-coefficient for the investor with habit preferences.

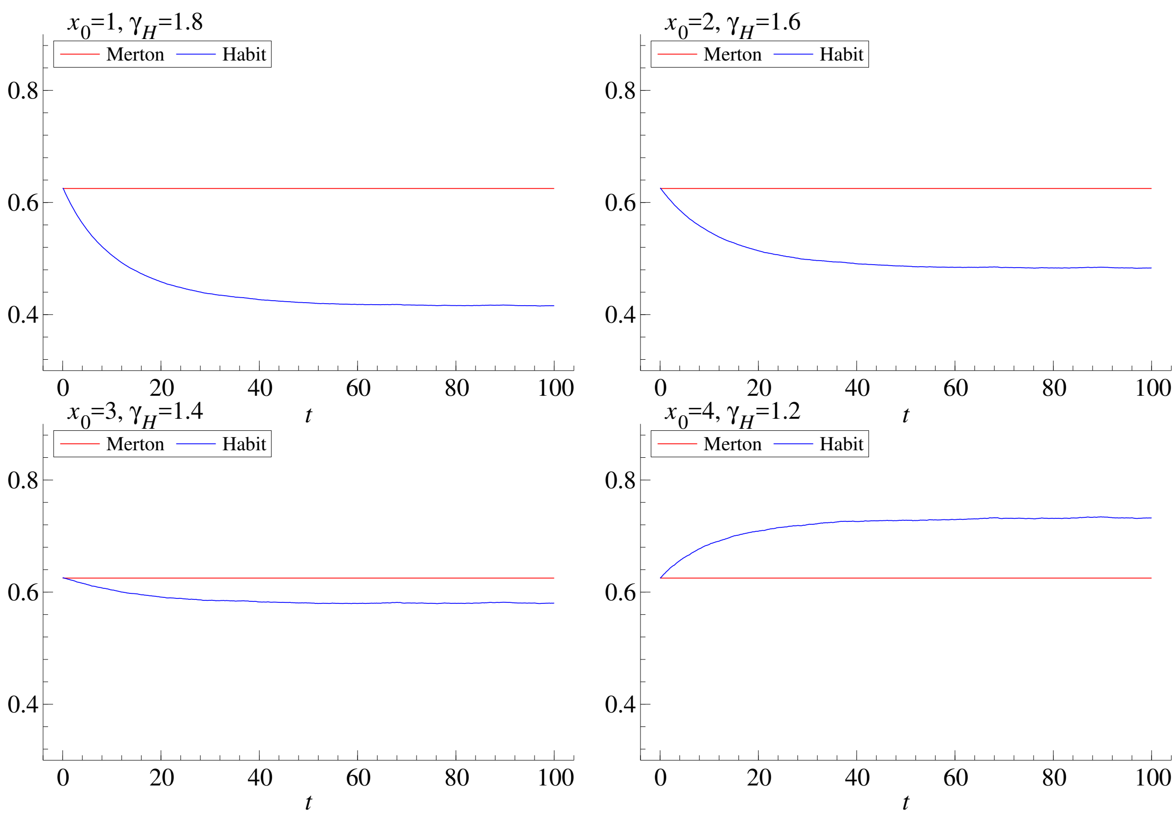

Figure 5: Average levels of the optimal risky portfolio share $\alpha(t)$ under habit formation. Parameter values are $\mu = 0.05$, $\sigma_z = 0.20$, $r = 0.05$, $\gamma = 2$, $a = 0.3$, $b = 0.25$, $\rho = 0.03$, and $W_0 = 100$. The averages are estimated from 10,000 simulated observations at each point in time. We use 10 time points per year. The parameter x_0 shows the initial habit level and γ_H shows the γ-coefficient for the investor with habit preferences.

structural, non-oil deficits. Although this structural deficit may not exceed 4% of the SWF's value, the actual deficit may be significantly larger. The rule does not explicitly say so, but the difference between the actual and the structural deficit must be financed by a special draw on the SWF because the Norwegian government does not, by statute, have any other source of financing as long as the value of the SWF stays above zero.

In addition, states often use fiscal policies actively as countercyclical measures. Although monetary policy arguably is superior in this regard, it may sometimes be unavailable, as is the case if the government seeks to maintain a fixed exchange rate or the country is a member of a monetary union. At the zero lower bound for interest rates, monetary policy typically loses some of its potency, thus leaving a case for fiscal policy. The Norwegian fiscal rule, as stated in the introduction, explicitly allows for discretionary fiscal policy to smooth the business cycle.

Countercyclical policy, including the automatic stabilizers, is thus one more way that the SWF can contribute to smoothing. Modeling the automatic stabilizers is equivalent to adding other income as an additional source to fund consumption over and above the return on investments. If markets are complete, the present value of this income can simply be added to the financial wealth without any further modification of the model. Typically, however, issues like moral hazard, dynamic inconsistency, and poor enforceability for contracts with sovereign governments mean that markets are not complete, so that a government's future tax revenues cannot be capitalized. As is well known, exact analytical solutions are then no longer available except in uninteresting cases like quadratic utility.

Discretionary fiscal policy can be modelled as an addition to the habit level driven by variations in exogenous income, so that the minimum level of consumption varies countercyclically. This naturally adds another layer of complexity to the model.

Two further elements of complexity are worth mentioning. One is that the return on risky assets is likely to be positively correlated with income. The other is that the risk-free rate is likely to move procyclically as well. All these complexities will thus have to be dealt with in a complete analysis of the investment and use of a SWF in a cyclical setting.

Viceira (2001) has worked out approximate solutions for optimal consumption and investment for an individual with stochastic labor income. However, he studies only permanent changes in labor income, arguing that transitory income variations are unimportant for individuals' saving and investment decisions. Permanent income changes may be important for governments as well, for example, regarding long-term pension obligations with defined benefits. Our focus, however, is on the cyclical income variations that give rise to automatic stabilizers and countercyclical fiscal policy, for which we believe the results will be somewhat different. Furthermore, Viceira's analysis ignores the complexities that are essential to our analysis, such as habit formation, time variation in the riskless rate, and discretionary, countercyclical fiscal policy. Viceira's assumption that all income changes are permanent is similar to the case of cointegration between stock and labor markets analyzed by Benzoni et al. (2007). However, Benzoni et al. (2007) also look at the case where cointegration breaks down in the late stages of a person's life, which may be more useful for our purposes.

Although conclusions from this work naturally are premature, we anticipate the following:

- Procyclical variations in other revenues create a new incentive for limiting financial risk taking. Although theory suggests effects going both ways, this seems to follow from the main finding by Moos (2011). We believe this effect will be stronger the more positive the correlation between financial returns and other revenues.

- Procyclical movements in the riskless rate should add to the demand for forward as well as backward smoothing, thus exacerbating the long-term volatility of consumption. This problem is special for governments with positive net financial assets because, for states with net indebtedness, financing becomes cheaper when interest rates are low.

- Long-term volatility will be even further exacerbated by the interaction between procyclical movements in income and financial returns as well as the countercyclical movements in habits implied by discretionary fiscal policy.

- All of the above effects are likely to exacerbate the time variation of the optimal investment strategy.

6 Conclusions

This paper has studied some of the issues that arise for a government that wants to fund part of its budget from the returns of a SWF. Although we have had to leave several important issues for future work, we feel entitled to make some conclusions. The most important one is that spending and investment decisions should not be made separately as the criteria for optimal investment generally depend on the way that the money is spent. In particular, if spending is to be smoothed in a backward as well as a forward sense, so that sharp spending changes are to be avoided, then the fund's portfolio should not simply be rebalanced after an unexpected drop in market values because a larger share then needs to be kept in secure assets so as to safeguard the financing of a smooth expenditure flow. Instead, the risky share of the portfolio should be reconsidered after each change in securities prices. Furthermore, the desire to maintain a smooth flow should generally motivate a conservative investment strategy.

We have also shown that the smoothing of current expenditure carries a price in terms of long-term volatility. That is, the greater the smoothing efforts for current expenditures, the greater the uncertainty will be regarding future consumption at longer horizons. The steady reconsideration of the investment strategy after each security price change helps limit this uncertainty, but only partially.

We furthermore believe that our analysis of consumption and investment with time variation in the riskless rate carries some new insight. It provides at least partial support for the claims that have been made in regard to the Norwegian SWF to the effect that the currently low real interest rates should temporarily indicate more modest draws on the fund. However, as long as these low rates can be considered temporary, our analysis indicates that the reduction can be more modest than would be appropriate if the current low levels represent a new long-term normal.

Our remaining conclusions follow more directly from the simplest models of spending and investment. Thus, we point out that the annual draw on a SWF depends on political preferences in a substantial way, meaning that favoring current generations over future ones can be perfectly rational, for example, if future generations can be expected to be more prosperous in general. Such preference may very well justify draws on the fund that exceed the normal rate of real return. The fund, and the absolute value of the annual draw, will then decline over time. However, this would then be the consequence of rational choices and neither irrational nor irresponsible.

However, if a government wants to use the fund's annuity value as the base for its annual draw (rather than a preference for current or future generations), it should make sure to use the risk-adjusted rate rather than a simple average. This insight has yet to find its way into the Norwegian fiscal rule.

References

Abel, A. B. (1990). Asset prices under habit formation and catching up with the joneses. *American Economic Review 80*(1), 38–42.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance 59*(4), 1481–1509.

Barro, R. J. (1979). On the determination of the public debt. *Journal of Political Economy 87*, 940–971.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein (2007). Portfolio choice over the life-cycle when the stock and labor markets are cointegrated. *Journal of Finance 62*(5), 2123–2167.

Campbell, J. and J. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy 107*(2), 205–251.

Campbell, J. and A. Deaton (1989). Why is consumption so smooth? *Review of Economic Studies 56*, 357–373.

Campbell, J. Y. and L. Viceira (2002). *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors* (first ed.). Oxford University Press.

Carroll, C., J. Overland, and D. Weil (2000). Saving and growth with habit formation. *American Economic Review 90*, 341–355.

Choudhry, M. (2007). *Bank Asset and Liability Management: Strategy, Trading, Analysis* (first ed.). Wiley Finance.

Constantinides, G. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy 98*, 519–543.

Di Maggio, M. and A. Kermani (2015). The importance of unemployment insurance as an automatic stabilizer. mimeo, Columbia Business School and University of California Berkeley.

Doornik, J. (1999). *Object-Oriented Matrix Programming Using Ox*. London: Timberlake Consultants Press and Oxford (www.nuff.ox.ac.uk/Users/Doornik).

Fama, E. F. and K. R. French (1988). Permanent and temporary components of stock prices. *Journal of Political Economy 96*, 246–273.

Flavin, M. (1985). The excessive sensitivity of consumption to current income - liquidity constraints or myopia? *Canadian Journal of Economics 1*(February), 117–136.

Giovannini, A. and P. Weil (1989). Risk aversion and intertemporal substitution in the capital asset pricing model. NBER Working Paper no. 2824.

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy 86*, 971–987.

King, M. and D. Low (2014). Measuring the ‘world’ real interest rate. NBER Working Paper 19887.

Lax, Y. (2002). Habit formation and lifetime portfolio selection. mimeo, Goldman Sachs Asset Management.

Matsen, E. (2003). Habit persistence and welfare gains from international asset trade. *Journal of International Money and Finance 22*(2), 239–260.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous time case. *Review of Economics and Statistics 51*, 247–257.

Moos, D. (2011). Portfolio and consumption decisions under mean-reverting returns and mean-reverting income growth. Ph.D. dissertation, University of St. Gallen.

NOU 2015:9. Finanspolitikk i en oljeøkonomi – praktisering av handlingsregelen (fiscal policy in an oil economy: The fiscal rule in practice). NOU 2015:9, Norwegian Ministry of Finance.

NOU 2016:20. Aksjeandelen i statens pensjonsfond utland (the equity share of the government pension fund global). NOU 2016:20, Norwegian Ministry of Finance.

Olsen, Ø. (2012). Economic perspectives. Address to the Supervisory Council of Norges Bank and invited guests.

Olsen, Ø. (2015). Economic perspectives. Address to the Supervisory Council of Norges Bank and invited guests.

Phelps, E. S. (1962). The accumulation of risky capital: A sequential utility analysis. *Econometrica 30*, 729–743.

Poterba, J. and L. Summers (1988). Mean reversion in stock returns: Evidence and implications. *Journal of Financial Economics 22*, 27–60.

Rachel, L. and T. Smith (2015). Secular drivers of the global real interest rate. IMF Staff Working Paper No. 571.

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics 51*(3), 239–246.

Summers, L. (2013). Crises yesterday and today. Remarks at the 14th Annual IMF Research Conference.

Sundaresan, S. (1989). Intertemporally dependent preferences and the volatility of consumption and wealth. *Review of Financial Studies 2*, 79–89.

Thwaites, G. (2014). Why are real interest rates so low? secular stagnation and the relative price of investment goods. Working Paper LSE and Bank of England.

Vasicek, O. A. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*(2), 177–188.

Viceira, L. M. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *Journal of Finance 56*(2), 433–470.

A Approximate solution of the model with time-varying risk-free rate

The problem is to maximize (1) subject to (6) and (7). Our method parallels that of Campbell and Viceira (2002), who studied this problem with long bonds as the only risky assets. We start with the Bellman equation:

$$0 = \max_{c(t),\alpha(t)} \left\{ e^{-\rho t} \frac{c(t)^{1-\gamma}}{1-\gamma} + \frac{1}{dt} E_t[dV(W(t), r(t), t)] \right\}. \tag{A.1}$$

After second-order expansion of the value function, application of Itô's lemma, and taking expectations, we can write this equation as:

$$\begin{aligned} 0 = \max_{c,\alpha} \Big\{ & e^{-\rho t} \frac{c^{1-\gamma}}{1-\gamma} + V_W[(r + \alpha\mu(r))W - c] + V_r\theta(r^* - r) + \frac{\partial V}{\partial t} \\ & + \frac{1}{2} V_{WW}\alpha^2 W^2 \sigma_z^2 + V_{Wr}\alpha W \beta \sigma_z^2 + \frac{1}{2} V_{rr}\sigma_r^2 \Big\}, \end{aligned} \tag{A.2}$$

where subscripts denote first and second-order partial derivatives, and we have, to simplify notation, omitted dating of the relevant variables.

Now, conjecture the following form for the value function:

$$V(W,r,t) = e^{-\rho t}\eta(r)^{-\gamma}\frac{W^{1-\gamma}}{1-\gamma}. \tag{A.3}$$

After substitution of the partial derivatives under this conjecture, the Bellman equation becomes:

$$\begin{aligned} 0 = \max_{c,\alpha} \Big\{ & e^{-\rho t}\frac{c^{1-\gamma}}{1-\gamma} + (1-\gamma)V[(r+\alpha\mu(r)) - c/W] \\ & - \gamma V\theta(r^* - r)\eta'/\eta - \rho V - \frac{1}{2}\gamma(1-\gamma)V\alpha^2\sigma_z^2 \\ & - \gamma(1-\gamma)V\alpha\beta(\eta'/\eta)\sigma_z^2 - \frac{1}{2}\gamma V[(\gamma+1)(\eta'/\eta)^2 - \eta''/\eta]\sigma_r^2 \Big\}. \end{aligned} \tag{A.4}$$

The optimization is now done conventionally by differentiating the expression inside the braces with respect to c and α, respectively, and equating these partial derivatives to zero. This gives the following first-order conditions:

$$c = \eta(r)W \tag{A.5}$$

and

$$\alpha = \frac{\mu(r)}{\gamma\sigma_z^2} - \beta\eta'/\eta. \tag{A.6}$$

Define $\mu^* = \mu(r^*)$ so that $\mu(r) = \mu^* - (1-\lambda)(r - r^*)$.

Substitution of these conditions into the Bellman equation gives us, after some algebra, the maximized value as:

$$\begin{aligned} 0 = & \gamma\eta + (1-\gamma)r - \gamma\theta(r^* - r)\eta'/\eta - \rho \\ & + (1-\gamma)\Big[\frac{\mu^*}{\gamma\sigma_z^2} - \beta\eta'/\eta - \frac{1-\lambda}{\gamma\sigma_z^2}(r - r^*)\Big][\mu^* - (1-\lambda)(r-r^*)] \\ & - \frac{1}{2}\gamma(1-\gamma)\Big[\frac{\mu^*}{\gamma\sigma_z^2} - \beta\eta'/\eta - \frac{1-\lambda}{\gamma\sigma_z^2}(r - r^*)\Big]^2\sigma_z^2 \\ & - \gamma(1-\gamma)\Big[\frac{\mu^*}{\gamma\sigma_z^2} - \beta\eta'/\eta - \frac{1-\lambda}{\gamma\sigma_z^2}(r - r^*)\Big]\beta\sigma_z^2\eta'/\eta \\ & - \frac{1}{2}\gamma[(\gamma+1)(\eta'/\eta)^2 - \eta''/\eta]\sigma_r^2. \end{aligned} \tag{A.7}$$

(A.7) is a second-order, non-homogeneous, non-linear ordinary differential equation in $\eta(r)$. Rather than trying to find an exact solution, however, we follow Campbell and Viceira in approximating the function $\eta(r)$ around r^* as follows:

$$\eta(r) \approx \eta^* + \eta^* \ln \eta(r - r^*),$$
$$\ln \eta(r - r^*) \approx C_0 + C_1(r - r^*).$$

We then easily find

$$\eta'/\eta \approx C_1,$$
$$\eta''/\eta \approx C_1^2.$$

Substituting this into (A.7) and ignoring quadratic terms in $(r - r^*)$, we get the following equation:

$$\begin{aligned} 0 =& \gamma\eta^*[1 + C_0 + C_1(r - r^*)] + (1 - \gamma)(r - r^*) + \gamma\theta(r - r^*)C_1 \\ &- (1 - \lambda)(1 - \gamma)\Big(\frac{\mu^*}{\gamma\sigma_z^2} - \beta C_1\Big)(r - r^*) + \text{terms not involving } r. \end{aligned} \tag{A.8}$$

This equation must hold as an identity in r. This means that we can use the method of undetermined coefficients to identify the parameters C_0 and C_1. In particular, the coefficients in (A.8) multiplying r must sum to zero. Noting that $\frac{\mu^*}{\gamma\sigma_z^2} - \beta C_1 = \alpha^*$, this gives the solution in the text.