Estimating Consumption Plans for Recursive Utility by Maximum Entropy Methods

Stephen Satchell, Susan Thorp and Oliver Williams
Estimating consumption plans for recursive utility by maximum entropy methods

Stephen Satchell
Trinity College, University of Cambridge, U.K.
and
Finance Discipline, University of Sydney, Australia

Susan Thorp
Finance Group, University of Technology, Sydney, Australia

Oliver Williams
Scalpel Research, 84 Brook Street, London, U.K.
and
Kings College, University of Cambridge, U.K.

January 28, 2012

Corresponding author: Susan Thorp, Finance Group, University of Technology Sydney, PO Box 123 Broadway NSW 2007 Australia; Tel: +61 2 95147784; email: susan.thorp@uts.edu.au
Abstract

We derive and estimate the optimal disbursement from an infinitely-lived charitable trust with an Epstein-Zin-Weil utility function, given general Markovian returns to wealth. We analyze two special cases: where spending is a power function of last period’s wealth and the endowment uses ‘payout smoothing’. Via nonlinear least squares, we estimate the optimal spending rate and the elasticity of intertemporal substitution for a trust with a typical diversified portfolio and for a portfolio of hedge funds. Finally, we use maximum entropy methods to characterize the returns distribution of a trust whose spending plan conforms with the optimality condition.

JEL classification: G23; D81; D91; E21

Key words: Intertemporal choice; Elasticity of intertemporal substitution; Moving average;

Endowed institutions, foundations and charitable trusts in the U.K. include universities, schools, research institutions, and grant-making charities. In 2012, the U.K. Charity Commission reported over 161,000 charities, holding investments in excess of £78 billion with annual spending over £53 billion. For the U.S., Standard & Poor’s Money Market Directories reported over 5,000 endowments and foundations, controlling more than $946 billion in assets. While there are a few studies of the U.S. and U.K. university endowment sectors, there is otherwise surprisingly little quantitative research published in this area, and the question of how best to spend the income and assets of an endowment remains a topic of interest and concern to trustees and regulators.

While many aspects of endowment spending policy are captured in core models of intertemporal optimization, there are important idiosyncrasies that warrant separate investigation. Here we extend existing results on optimal spending plans for infinitely-lived entities by incorporating recursive preferences, predictable returns and payout smoothing policies. Payout smoothing, or, the use of averages of past and current wealth as the base for current expenditures, has not been analyzed in earlier theoretical work but is a key feature of observed endowment behavior. We present implicit analytical solutions to this problem in a general setting for returns. In particular, we relate the payout smoothing consumption spending rule to the dynamic structure of the returns process, resulting in an endogenously generated non-linear returns dynamics.

We apply two approaches to characterizing the consumption Euler equation for general returns processes. First we assume lognormally distributed errors for the implied moment conditions and estimate both the optimal disbursement rate and the elasticity of intertemporal substitution (EIS) of a representative charity, using non-linear methods applied to historical endowment returns. At
discount rates of 3% p.a., optimal disbursement rates of a typical U.K. charity are 2.6% of wealth p.a. in real terms and the elasticity of intertemporal substitution is 1.3.\(^4\)

Next we relax the assumption of lognormal errors and introduce maximum entropy methods to examine how the characteristics of equilibrium returns distributions relate to the consumption spending rule. The consumption Euler equation places constraints on serial dependence in the returns distribution, conditional on preference parameters, so we aim to find stationary returns distributions consistent with these constraints. This time we fix discount and disbursement rates then numerically estimate maximum entropy returns distributions using historical data from two representative endowment portfolios. The result is a complete characterization of stationary equilibrium returns distributions, including an estimate of serial dependence, for a matrix of risk and EIS parameters. For example, at risk aversion and intertemporal substitution parameter values above one, we estimate serial correlation in monthly (fund of hedge fund) returns close to 0.3, with good fit of the maximum entropy distribution according to a Kolmogorov-Smirnov test.

We set our model up with several distinct features of charities in mind. First, since many charitable endowments are constrained by charter to provide funding perpetually to clients or projects with short time horizons, we define the problem as choosing annual spending rates over an infinite horizon. Secondly, Brown et al. (2010a) and Acharya and Dimson (2007) note that university endowment structures often decentralize investment management. Endowment boards or investment committees make high level investment policy but day to day decisions are often delegated to groups of fund managers or to sub-committees. As a result, asset allocation decisions are subject to general investment objectives (and sometimes to benchmark settings) that are informed by long-term spending requirements, but are made separately from consumption plans. Consequently, we derive spending plans conditional on a pre-set investment allocation. Thirdly, charities and endowments frequently allocate large proportions of their portfolios to alternative asset classes (Brown et al. 2010b) and log portfolio returns are unlikely to be normally distributed, so we extend our analysis to more general returns distributions. Fourthly, charitable trusts face volatile returns, but make disbursements to beneficiaries who often value smooth funding streams. Models of the disbursement rate of charities that apply the usual time-separable expected utility functions overlook the fact that, for these organizations, risk aversion and aversion to intertemporal substitution appear conceptually and practically distinct. Charities tolerate considerable uncertainty over returns
while aiming for fairly smooth payments to beneficiaries over time. We work with the recursive or non-expected utility preferences proposed by Epstein and Zin (1989) and Weil (1990), which allow a partial separation of tastes for risk and intertemporal consumption. Finally, Brown et al. (2010a) note that endowments and charities typically employ a ‘payout-smoothing’ model where a percentage is applied to a multi-year moving average of past endowment values. However \textit{ex post} observed spending patterns of U.S. university endowments still deviate from this mechanical rule, showing sensitivity to contemporaneous negative wealth shocks. Endowments that follow the smoothing rule endogenize predictability into the returns to wealth and we model these properties.

In the next section we set out the recursive utility model and existing results for independent and identically distributed returns (i.i.d.). Section 2 then analyzes spending rules for three increasingly general patterns of non-i.i.d. returns and presents estimates of preference parameters for a typical charitable trust. In section 3 the maximum entropy method is discussed and implemented. Section 4 concludes.

1 Model

Consider an entity that is infinitely lived but makes annual consumption plans. Giovannini and Weil (1989) and Weil (1990) write the discrete-time recursive utility function for such an entity, that aggregates current consumption and uncertain future consumption, as:

\[
L_t = U[c_t, E_tL_{t+1}] = \left\{ \frac{(1 - \delta)c_t^{1-\rho} + \delta[1 + (1 - \delta)(1 - \alpha)E_tL_{t+1}]^{\frac{1-\alpha}{1+\rho}}}{(1 - \delta)(1 - \alpha)} \right\}^{\frac{1}{1-\rho}} - 1
\]

where \( \delta \in (0, 1) \), \( \alpha > 0 \) and \( \rho > 0 \), and where \( C_t \) is payments to worthy causes and costs.

Convexity (\( \alpha > \rho \)) implies more rapidly increasing patience, and concavity (\( \alpha < \rho \)), more slowly increasing patience, as expected future utility rises. Entities which are more risk tolerant and value smoothness (\( \alpha < \rho \)) prefer late resolution of uncertainty, and entities who dislike risk but tolerate larger swings in certainty equivalent utility (\( \alpha > \rho \)), prefer early resolution. Under the special case where \( \alpha = \rho \), the utility function represents the preferences of an individual with constant relative risk-aversion (CRRA). In the CRRA case, the discount parameter \( \delta \) is a direct measure of
impatience as the inverse of (one plus) the rate of time preference whereas in the recursive utility case time preference is a more complex function of model parameters, including $\delta$.

Another way to view the parameters of the model is to recognize that the coefficient of relative risk aversion for timeless gambles is $\alpha$ and the constant elasticity of intertemporal substitution for deterministic consumption paths is $\frac{1}{\rho}$. If either parameter approaches unity, then preferences become logarithmic in that dimension, resulting in logarithmic risk preferences when $\alpha \to 1$ and logarithmic intertemporal substitution preferences when $\rho \to 1$. Under the special (CRRA) case where $\alpha = \rho$, the inverse of the risk aversion parameter is the elasticity of intertemporal substitution.

1.1 Wealth

Setting aside questions of portfolio allocation, and assuming that no donation income is received, the budget constraint of the charitable endowment is

$$w_{t+1} = (w_t - c_t)Z_t$$

(2)

where $w_t$ is wealth at time $t$ and $Z_t$ is the random growth in investments from $t$ to $t + 1$. If $c_t = A(Z_{t-1})w_t$, where $A(Z_{t-1})$ is a general expression for the optimal spending rule as a function of all past realizations of $Z$ (not only the current realization $Z_{t-1}$), then equation (2) is

$$w_{t+1} = [1 - A(Z_{t-1})]w_tZ_t.$$  

(3)

The difference equation in wealth can be written as,

$$w_t = w_0 Y_{t-1} \prod_{i=0}^{t-1} [1 - A(Z_i)]$$

(4)

where $Y_{t-1}$ is the accumulated value of one unit of wealth invested at $t = 0$ and held until time $t$, which we assume is random and non-negative but otherwise unrestricted.

However, if $Z_i$ is a positive, independent and identically distributed (i.i.d.) random variable, and $Z_i^{1-\alpha}$ is a well defined random variable such that $E(Z_i^{1-\alpha}) = \phi$ exists for $0 < \alpha < \infty$, it follows that $E(Y_t^{1-\alpha}) = \phi^t$ for all integer $t > 0$. Further, in this special case, Giovannini and Weil (1989)
and Weil (1990) derive a constant proportional spending rule where:

\[ \bar{A} = 1 - \left( \delta \phi \right)^{\frac{1}{\rho}}. \]

While returns to endowment portfolios may be i.i.d., several features of endowment investment practice and spending patterns weigh against this simplification. First, many endowments allocate substantial fractions of capital to alternative asset classes such as hedge funds, private equity and infrastructure (Brown et al. 2010b) which exhibit correlated returns (Getmansky et al. 2004). Secondly, the multi-period average spending policy itself will generate correlation in the returns to wealth since past wealth levels will predict future wealth levels.

In what follows we assume an asset allocation decided by computation or committee, that may or may not be optimal, and calculate the the conditionally optimal disbursement rate for increasingly more general cases that are of practical interest to endowment managers. First, we specify a power rule where the disbursement rate is a constant scaling of a power function of current (not necessarily i.i.d.) investment returns then analyze the case where the spending rate is a constant scaling of a moving average of wealth, reflecting the multi-period smoothing practices of many endowments, and third where returns to wealth follow general Markovian processes.  

2 Spending rules for general wealth processes

Begin with the Euler equation in consumption from equation (1) set out in Giovannini and Weil (1989) as:

\[ \delta^{\frac{1-\alpha}{1-\rho}} E_t \left[ \frac{c_{t+1}}{c_t} \left( \frac{1}{\rho} \right) \left( \frac{1-\alpha}{1-\rho} \right) Z_t^{\frac{1-\alpha}{1-\rho}} \right] = 1 \]  

Recall that \( A(Z_{t-1}) \) is an optimal spending rule that is a general function of past and current realizations of wealth, so that

\[ c_t = A(Z_{t-1}) w_t, \]
and

\[
\frac{c_{t+1}}{c_t} = \frac{A(Z_t)w_{t+1}}{A(Z_{t-1})w_t} = \frac{A(Z_t)Z_t(1 - A(Z_{t-1}))}{A(Z_{t-1})}.
\]

To simplify notation, define \( \theta = \frac{1 - \alpha}{1 - \rho} \) so (5) becomes

\[
\delta^\theta E_t \left[ (A(Z_t)Z_t(1 - A(Z_{t-1})))^{\rho \theta} Z_t^\theta \right] = A(Z_{t-1})^{-\rho \theta}
\]

or

\[
E_t \left[ (A(Z_t)Z_t(1 - A(Z_{t-1})))^{\rho \theta} Z_t^\theta \right] = A(Z_{t-1})^{-\rho \theta} \delta^{-\theta}.
\]

Under i.i.d. returns, Giovaninni and Weil (1989) show that spending plans depend on \( \rho \) rather than \( \alpha \), a result we generalize in the following proposition for the Euler equation (8). We label the rule in (8) as the ‘smoothing-consistent dynamic process’ and we define \( V_t \) as a stochastic process where \( E_t(V_t^\theta | Z_{t-1}) = 1 \). This can be viewed as some standardization of \( Z_t \) that sets \( \theta \) to zero.)

**Proposition 1** Assuming that equations (5) and (6) are satisfied and that \( V_t \) is as defined above, the smoothing-consistent dynamic process has dynamics that do not depend upon risk aversion \( \alpha \), but do depend upon the elasticity of intertemporal substitution \( 1/\rho \).

**Proof.** Equation (8) can be written as

\[
\delta^\theta E_t \{ [A(Z_t)]^{-\rho \theta} Z_t^{\theta(1-\rho)} \} = [A(Z_{t-1})]^{-\rho \theta} [1 - A(Z_{t-1})]^{\rho \theta}
\]

therefore

\[
[A(Z_t)]^{-\rho \theta} Z_t^{\theta(1-\rho)} = \delta^{-\theta} [A(Z_{t-1})]^{-\rho \theta} [1 - A(Z_{t-1})]^{\rho \theta} V_t^\theta.
\]

It follows that

\[
[A(Z_t)]^{-\rho} Z_t^{(1-\rho)} = \delta^{-1} [A(Z_{t-1})]^{-\rho} [1 - A(Z_{t-1})]^{\rho} V_t.
\]
Since $1/\rho$ is the crucial parameter in this case, as for the i.i.d. case, we focus below on its estimation and that of the spending rate rather than other preference features such as risk aversion.

### 2.1 Power spending rule

We now explore two important special cases of (8) with the aim of deriving expressions that allow estimation. If we restrict $A(Z_t)$ to a power form where, $A(Z_t) = AZ_t^b$, after some re-arrangement, we obtain

$$E_t \left[ Z_t^{\theta(1-\rho(b+1))} \right] = \left( \frac{Z_{t-1}^b}{1 - AZ_{t-1}^b} \right)^{-\rho \theta} \delta^{-\theta}. \quad (9)$$

Using (9) and assuming a multiplicative error process for $Z_t$ such that $E_t(V_t^{\theta(1-\rho(b+1))} \mid Z_{t-1}) = 1$, this implicit expression for $A(Z_t)$ can be rearranged for estimation:

$$Z_t^{\theta(1-\rho(b+1))} = \left( \frac{Z_{t-1}^b}{1 - AZ_{t-1}^b} \right)^{-\rho \theta} \delta^{-\theta} V_t^{\theta(1-\rho(b+1))} \quad (10)$$

$$Z_t = \left( \frac{Z_{t-1}^b}{1 - AZ_{t-1}^b} \right)^{-\frac{\theta}{1-\rho(b+1)}} \delta^{-\frac{1}{1-\rho(b+1)}} V_t \quad (11)$$

$$\log Z_t = d + D_1 \log \left( 1 - AZ_{t-1}^b \right) - D_1 \log Z_{t-1} + \log V_t. \quad (12)$$

allowing for estimation of $d$, $D_1 = \frac{\rho}{1-\rho(b+1)}$, $A$ and $b$ and the parameters of error process $V_t$. In this way, by unpacking the expectation in equation (9), we arrive at a dynamic non-linear time series process, equation (10), consistent with the spending rule in equation (6).

The original parameters can then be recovered using the relationships:

$$\rho = \frac{D_1}{1 + D_1(b + 1)} \quad (13)$$

and

$$\log \delta = -\frac{d}{1 + D_1(b + 1)} \quad (14)$$

If we assume that $0 < b < 1$ and impose the constraint $|D_1 b| < 1$ then this leads to the following
requirements:

\[ d > 0 \]  \hspace{1cm} (15)  
\[ D_1 > 0 \]  \hspace{1cm} (16)  

which ensure that \( 0 < \delta < 1 \) and \( \rho > 0 \) respectively. In addition, to ensure that the whole of wealth is not consumed in a single period, we need

\[ AZ_t^b < 1 \quad \forall t. \]  \hspace{1cm} (17)  

Since \( Z_t \) is a gross portfolio return, this will ‘usually’ be easily satisfied if \( 0 < b < 1 \) and the consumption proportion is plausible percentage of total wealth (such as \( A \approx 0.03/12 \) for monthly data), except in the case of an improbably large investment return. 6

### 2.2 Non-Markovian spending rule from smoothed endowment value: a ‘real-world’ example

We now look into the case where the endowment consumes a fixed proportion of wealth averaged over the past 36 months. Brown, et al. (2010a, p.4) note that ‘...the vast majority of [endowment spending] policies use a multi-year, moving average of past endowment values as the basis to which the payout percentage applies’. Since the returns data used for estimation below are monthly, we define the disbursement rate as a proportion of the 36-month-average (monthly) return,

\[ A(Z_{t-1}) = a \left( \prod_{j=0}^{m-1} Z_{t-1-j} \right)^\frac{1}{m} \]  \hspace{1cm} (18)  

where \( m = 36 \) and \( a > 0 \) is a constant. In this case we can write:

\[ \frac{A(Z_t)}{A(Z_{t-1})} = \left( \frac{Z_t}{Z_{t-m}} \right)^\frac{1}{m} \]  \hspace{1cm} (19)  

\[ (1 - A(Z_{t-1}))^{-\rho \theta} = \left( 1 - a \left( \prod_{j=0}^{m-1} Z_{t-1-j} \right)^\frac{1}{m} \right)^{-\rho \theta} \]  \hspace{1cm} (20)  

9
and using these expressions, we can rewrite (8), to obtain an implicit expression for the multi-period spending policy:

\[
E_t \left[ \left( \frac{Z_t}{Z_{t-m}} \right)^{-\rho \theta \over m} Z_t^{\theta (1-\rho)} \left( 1 - a \left( \prod_{j=0}^{m-1} Z_{t-1-j} \right) \right) \right] = \delta^{-\theta}. \tag{21}
\]

If we again assume a multiplicative error for \(Z_{t-1}\), such that \(E_t(V_{t}^{-\rho \theta \over m} + \theta (1-\rho) | Z_{t-1}) = 1\) then we can write,

\[
Z_{t}^{-\rho \theta + \theta (1-\rho)} = \left( 1 - a \left( \prod_{j=0}^{m-1} Z_{t-1-j} \right) \right) Z_{t-m}^{\rho \theta \delta^{-\theta} V_{t}^{-\rho \theta + \theta (1-\rho)}}. \tag{22}
\]

This can be further simplified by recalling that \(\theta = \frac{1-\alpha}{1-\rho}\), which gives

\[
Z_{t}^{-\rho (1+m) \over m(1-\rho)} = \left( 1 - a \left( \prod_{j=0}^{m-1} Z_{t-1-j} \right) \right) Z_{t-m}^{\rho \delta \over 1-\rho} V_{t}^{-(1-\rho) \rho (1+m) \over m(1-\rho)}. \tag{23}
\]

Now defining \(f = \rho (1 + m) - m\), and taking logs,

\[
\log Z_t = -{m \rho \over f} \log \left( 1 - a \left( \prod_{j=0}^{m-1} Z_{t-1-j} \right) \right) + \rho \log Z_{t-m} + {m \over f} \log \delta + \log V_t \tag{24}
\]

If there is no averaging \((m = 1)\) then (24) simplifies to (12) in the special case of \(b = 1\).

### 2.3 Existence of a stationary spending rule under Markovian returns

Analytical results so far have not relied on strict restrictions on the returns process. However if log returns are Markovian, the process can be expressed as \(\phi(\log Z_t) = \psi(\log Z_{t-1}) + \log \varepsilon_t\), and \(\log \varepsilon_t \sim i.i.d.(0, \sigma^2)\), is independent of \(\log Z_{t-1}\). For the case of multiple lags, such as the MA(36), log returns can be expressed as a vector Markovian process. (See Anderson (1971), p177.)

For the special case analogous to the power spending rule in (12) where \(\log Z_t = \psi(\log Z_{t-1}) + \log \varepsilon_t\), we can write \(h(\cdot)\) for the stationary \(pdf\) of \(y = \log Z_t\) and \(g(\cdot)\) for the \(pdf\) of the error term.
\[ V_t. \text{ If } x = \log(Z_{t-1}), \text{ then} \]

\[ h(y) = \int_{-\infty}^{\infty} h(x)g[y + \psi(x)]dx. \quad (25) \]

Conditions for the existence of such an \( h(\cdot) \) given \( g(\cdot) \) are discussed in Tong (1983, chapter 4). Verification and analytical calculation of these existence conditions are beyond the scope of this paper but in section 3 below we estimate stationary distributions for the equilibrium spending rule.\(^7\)

### 2.4 Empirical results

We use two series of investment returns to illustrate the empirical implications.\(^8\) First, we simulate returns to the pre-existing portfolio structure of an independent UK biomedical research-funding charity, the Wellcome Trust (Wellcome 2005).\(^9\) Monthly (log) real portfolio returns run from January 1990 to July 2010, (247 observations) where individual asset class returns are taken from standard indexes, and deflated using consumer prices and earnings data. Second, we take monthly (log) real net-of-fees returns to the Hedge Fund Research Composite Index, deflated using the same method. (Appendix details all data sources and calculations.) Summary statistics in Table 1 show that both series are negatively skewed and leptokurtic with significant serial correlation.

<table>
<thead>
<tr>
<th>Mean (^a)</th>
<th>Volatility</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>AR(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.05%</td>
<td>13.30%</td>
<td>-0.73</td>
<td>3.92</td>
<td>0.136 (^b)</td>
</tr>
<tr>
<td>7.90%</td>
<td>7.16%</td>
<td>-0.79</td>
<td>5.651</td>
<td>0.274 (^c)</td>
</tr>
</tbody>
</table>

\(^a\) Annualized monthly returns
\(^b\) \(p < 0.05\)
\(^c\) \(p < 0.01\)

Table 2 shows results of non-linear least squares estimation of equation (24) where \( \delta \) is fixed at 0.9975 (3% annual discount rate), \( m = 36 \) and \( \rho \) is constrained to be less than two. We estimate a consumption rate is 0.2% per month, or 2.6% p.a. The estimates are by non-linear least squares using a Gauss-Newton method with starting points determined by grid search. Standard errors are
computed by linear approximation, as the expectation of the inverse of the outer product of the gradient vector of the likelihood function.

Table 2: Estimated disbursement rate and inverse elasticity of intertemporal substitution

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>0.744</td>
<td>0.142</td>
<td>5.251</td>
<td>0.000</td>
</tr>
<tr>
<td>$a$</td>
<td>0.0022</td>
<td>0.0009</td>
<td>2.508</td>
<td>0.013</td>
</tr>
</tbody>
</table>

$^\delta = 0.9975$; Residual standard error = 0.0381, 209 d.f.

3 Determining the Returns Distribution by Maximum Entropy

Our derivation of smoothing-consistent dynamic processes in Section 2 takes an Euler equation as its point of departure. This Euler equation places restrictions on the moments of the returns process $Z_t$ but is silent as to the exact form of the stationary distribution of returns or the associated error process. Nevertheless we have also shown that further moment conditions can be imposed (e.g. $E[V_t^\theta | Z_{t-1}] = 1$) so that estimation and inference can be carried out under plausible distributional assumptions, in our case, normality in the log error process.

The functional forms of the spending rules which we have considered so far have been relatively simple and hence a fairly straightforward manipulation such as the log transform has been enough to obtain an estimable equation. However we have drawn attention to the need to ensure the boundedness of the spending rule, that is, $0 < A(Z_t) < 1$, which introduces a potentially awkward non-linearity. Loosely speaking, our problem has three ‘unknowns’: (a) a spending rule (which must be bounded), (b) a returns process (which we might wish to be ergodic/stationary), (c) an error distribution. We can place various restrictions on any two of these and estimate the parameters of the third and in the foregoing sections we have predominantly focused on (a) and (b).

We discuss next an alternative approach where we focus instead on the bivariate finite-dimensional distribution of $(Z_{t-1}, Z_t)$ as an alternative to examining (b) and (c) separately. Although we will not explicitly derive the structure of the returns process, it is well known that the Kolmogorov existence theorem provides consistency conditions under which a family of finite-dimensional dis-
tributions defines a stochastic process (see, for instance, Billingsley (1995)) and our approach in this section could be extended with that purpose in mind.

For the purposes of exposition we focus here on a spending rule which depends only on current period returns, however it is straightforward to extend our treatment to cases involving smoothing. We begin with the Euler equation (8)

\[ E_t \left[ (A(Z_t)Z_t(1 - A(Z_{t-1})))^{-\theta} Z_t^\theta \right] = A(Z_{t-1})^{-\theta} \delta^{-\theta} \]

and apply the law of iterated expectations to obtain a bivariate moment condition in terms of the two random variables \(Z_{t-1}\) and \(Z_t\) as follows:

\[ E \left[ \left( \frac{A(Z_t)}{A(Z_{t-1})}(1 - A(Z_{t-1})) \right)^{-\theta} Z_t^{1-\alpha} \right] = \delta^{-\theta} \]

We will now determine a probability distribution of \((Z_{t-1}, Z_t)\) consistent with this moment condition along with any other constraints which may be relevant such as unconditional means and variance of returns. There is an extensive treatment of this moment problem in mathematics (see, for example, Akhiezer 1965; Ang et al. 2002) which, from a theoretical perspective, deals with questions such as existence and uniqueness of distributions given a sequence of moments. This is complemented by a broad empirical literature which offers various alternative methods of distribution construction based on such techniques as Padé approximations and expansion in orthogonal polynomials. However a popular approach which has become common across many scientific fields is the method of maximum entropy, originally proposed by Jaynes (1957). For a continuously-distributed random vector \(Z\), the fundamental method is to find the probability density \(p(Z)\) which maximizes the quantity (entropy) \(-\int_R p(Z) \log p(Z) d\nu\) subject to the constraints that various moments of interest must equal their required values, where \(R\) represents the support of the distribution.

For our example, we choose to use five moment conditions (in addition to the obvious condition that the density must integrate to unity). These are: two mean conditions (we require that the mean of \(Z_{t-1}\) and \(Z_t\) both equal our sample means), two variance conditions (expressed as constraints on \(Z_{t-1}^2\) and \(Z_t^2\); again we require that variances equal the sample variance) and finally the Euler condition (8) above. A particularly appealing aspect of this approach is that we can explicitly
incorporate a bounded spending function into the Euler equation and in this case we choose to use
\[ A(Z_{t-1}) = \frac{cZ_{t-1}}{1+cZ_{t-1}}. \]

It is straightforward to show that the density which solves this problem takes the form \( p(Z) = \exp\left[-\sum_{i=0}^{5} \lambda_i f^{(i)}(Z)\right] \) where \( Z \equiv (Z_{t-1}, Z_t) \), \( \lambda_i \) is the Lagrange multiplier associated with the \( i \)'th moment condition and \( f^{(i)} \) represents the function underlying the \( i \)'th moment, that is, we have:

\[
\begin{align*}
  f^{(0)} &= 1 \\
  f^{(1)} &= Z_t \\
  f^{(2)} &= Z_{t-1} \\
  f^{(3)} &= Z_t^2 \\
  f^{(4)} &= Z_{t-1}^2 \\
  f^{(5)} &= \left( \frac{A(Z_t)(1-A(Z_{t-1}))}{A(Z_{t-1})} \right)^{-\rho} Z_t^{1-\alpha} \quad \text{with} \quad A(Z_{t-1}) = \frac{cZ_{t-1}}{1+cZ_{t-1}}.
\end{align*}
\]

Given the functions above it is clear (from inspection of the functional form of this density) that we may well obtain a density which somewhat resembles the bivariate normal, but instead of a conventional covariance term (in the product \( Z_{t-1}Z_t \)) we have a more complex characterization of dependency via the term in \( f^{(5)}(Z_{t-1}, Z_t) \). We note also that the moment condition which we applied to the multiplicative error term in section 2.2 relates specially to the particular formulation of the \( Z_t \) process which we outlined in that section, and hence we do not deal with that condition here since we approach the estimation problem from a different perspective as previously discussed.

Unfortunately the analytical simplicity of this result is counterbalanced by some numerical difficulties which are involved in obtaining the values of \( \lambda_i \). Various early algorithms were proposed by Johnson (1979) and Agmon et al. (1979) among others, and Zellner and Highfield (1988) presented a numerical methodology which they applied to an example of a distribution with four moments, the Cobb-Koppstein-Chen (1983) family. However in some circumstances it can be challenging to achieve convergence by the Newton-based methods typically used in this literature and Ormoneit and White (1999) subsequently proposed improvements to the basic algorithm to ensure convergence; further recent advances have been presented by Rockinger and Jondeau (2002), Wu (2003) (a sequential updating method) and Chen, Hu and Zhu (2010) (a hybrid method of linear
Figure 1: Relationship between $\alpha$ and $\rho$ implied by applying method-of-moments to sample returns data, assuming $c = 0.0022$ and $\delta = 0.9975$. (Solid line is the charity portfolio, dashed line is the hedge-fund index.)

Clearly the particular values of $\lambda_i$ will depend on our model parameters $(\alpha, c, \delta, \rho)$. We are therefore faced with an identification challenge since these four parameters all feature together in only one moment condition ($f^{(5)}$). To address this we first fix $\delta = 0.9975$ and $c = 0.0022$ (inspired by the results of empirical analysis in Section 2.4, although our analysis here is entirely separate), then to investigate plausible values of $(\alpha, \rho)$ we apply the conventional method-of-moments to the sample data and in Figure 1 we plot $(\alpha, \rho)$ pairs which are consistent with the data, assuming a limited range of $\alpha$ values.\(^{12}\)

In Figures 2 and 3 we present our numerically-obtained maximum entropy densities for $(Z_{t-1}, Z_t)$ for various example parameter combinations for the hedge fund and charity datasets respectively. We have fixed $\alpha$ at various levels and used values of $\rho$ close to those suggested by the method-of-moments (as illustrated in Figure 1) and we emphasize that these $(\alpha, \rho)$ values have been chosen for illustrative purposes and have not been determined by an optimization process.\(^{13}\) In each case we also provide the Pearson correlation coefficient calculated for that particular distribution.\(^{14}\)

It is apparent that the relationship between $(\alpha, \rho)$ and correlation is non-obvious, in other words, over the range of risk aversion which we have chosen it tends to be the case that increasing $\rho$ is associated with increasingly positive correlation in returns, however more general quantitative insights are hard to come by, especially if parameter values are examined over a larger range.

Nevertheless an important strength of our maximum entropy approach is that it does enable us to
translate the relatively opaque Euler condition \( f^{(5)} \) into a correlation value in this fashion, which provides a basis for further analysis on more familiar ground. Evidently our examples include several plausible parameter combinations which would be consistent with trending in returns, as well as combinations associated with negative first-order autocorrelation.

We now investigate how well our maximum entropy distributions fit our sample datasets. In principle we could proceed to estimate parameters by maximum-likelihood, subject to imposing identification restrictions, but since the maximum entropy density is not parameterized directly in terms of the economic parameters (such as \( \alpha \)) this would require that the entire process of density determination be repeated with each iteration of the maximum likelihood procedure, with fresh values of \( \lambda_i \) being solved each time, dependent on model parameters at that particular step. Furthermore we would need to establish an expression for the joint density of all sample observations, not just the pair \((Z_t, Z_{t+1})\). Since that effort is largely of numerical interest (and beyond the scope of this paper) we have not pursued it and we instead take a goodness-of-fit approach. Our method is to compute Kolmogorov-Smirnov (KS) statistics and accompanying \(p\)-values which relate to two-dimensional goodness-of-fit tests between our example densities and the sample datasets. This is calculated according to the method of Fasano and Franceschini (1987) using the algorithm of Press et al. (2007).\(^{15}\) Our KS results are incorporated into the captions in Figures 2 and 3 and clearly there are several \((\alpha, c, \delta, \rho)\) combinations where we are unable to reject the hypothesis that returns were generated by our entropy-maximizing distributions.\(^{16}\)

Hence the results which we have presented prove by construction that a stationary distribution for \( Z_t \) can exist which is consistent with recursive utility and serially-correlated returns, as well as a particular form of consumption rule. Having determined such a distribution, our earlier stochastic dominance results have immediate applicability. Furthermore the maximum entropy method of distribution construction is a useful way of shedding light on the correlation induced in equilibrium by a particular spending rule and profile of structural parameters.
Figure 2: Maximum entropy densities of hedge fund returns ($Z_t - Z_{t-1}$) for various $(\alpha, \rho)$ combinations; $\delta = 0.9975$ and $c = 0.0022$ in all cases. KS indicates Kolmogorov-Smirnov statistic with $p$-value in parentheses.
Figure 3. Maximum entropy densities of charity portfolio \((Z_{t-1}, Z_t)\) for various \((\alpha, \rho)\) combinations; \(\delta = 0.9975\) and \(c = 0.0022\) in all cases. KS indicates Kolmogorov-Smirnov statistic with \(p\)-value in parentheses.
4 Conclusion

We address the problem of choosing an optimal disbursement rate from a charitable trust or endowment with EZW preferences. Since endowments invest in a diverse range of asset classes, and also commonly use a multiperiod average spending rule, the problem requires a solution that allows for a general returns process, not necessarily log-normal, and that accommodates the inherent serial correlation induced by spending rules based on averaged wealth. Our analysis addresses both of these complications, deriving an implicit and general power spending rule from Euler equations. Consistent with existing results for an i.i.d. wealth process, spending plans depend on the size of the EIS rather than the risk preference parameter. Further, using non-linear methods and data from a representative endowment portfolio, we estimate values for the ideal disbursement rate and the elasticity of intertemporal substitution, which is a key parameter in spending rules under recursive preferences.

The Euler equation governing optimal spending plans implies constraints over the equilibrium returns distribution that we analyze using maximum entropy methods. Estimation using two representative datasets demonstrates the existence of stationary distributions consistent with recursive utility and serially correlated returns for the specific consumption rules derived here. While we have considered the plans of infinitely-lived endowments, these results can be applied to other problems in recursive utility maximization with serially correlated state variable dynamics.
Notes


4In practice, some charities may be restricted by regulation to minimum disbursement quotas (rates of spending out of accumulated wealth) that could reduce the welfare of the trust. The Canada Revenue Agency, for example, currently requires that 3.5% of value of property owned by a charity (averaged over the two years prior to the current fiscal year) but not used directly in activities or administration, be disbursed each year. Our analysis suggests that for some preference patterns and empirical settings, such regulations may be a binding constraint which reduces the welfare of the charitable trust. We thank Mr Vincent Taubman of TD Asset Management for advice on this issue. See http://www.cra-arc.gc.ca/E/pub/tg/t4033-1/t4033-1-10e.pdf

5Bhamra and Uppal (2006) set out the implicit portfolio optimality condition, and explicit optimal portfolio weights under a two-state process for the risky asset.

6Alternatively, the range of log(\(V_t\)) could be constrained.

7Tong demonstrates such a calculation in exercise 4.7, p140

8The general conditions for consistent estimation of the parameters of Euler equations such as [5] are discussed in Hansen (1982), Eichenbaum, Hansen and Singleton (1987) and Singleton (1988) among others, requiring stationarity of the ratios of random variables (such as current and future consumption) while the random variables themselves may exhibit stochastic trends.

9The Wellcome Trust Annual Report states their principal investment objective as ‘total return in inflation-adjusted terms over the long term in order to provide for real increases in annual expenditure while preserving at least the Trust’s capital base in real terms’. This matches up with the investment objectives of the majority of Oxbidge College endowments. Acharya and Dimson (2007) report that more than 60% of endowments in this group choose ‘Maximize long-term total return at an acceptable level of risk’ or ‘Long-term preservation of capital with a reasonable and predictable level of income’ as their main objective.

10Loosely speaking this amounts to solving for the distribution which can be realized in the maximum number of ways (in terms of possible elemental outcomes) while being consistent with the moment conditions. This ‘maximum entropy’ distribution is often informally described as being the ‘smoothest’ choice, or the choice which makes the minimal assumptions necessary to achieve compliance with the moment conditions. Cover and Thomas (2006) provide a thorough explanation of the approach and Joe (1997) demonstrates its relevance in the multivariate context. A particular strand of literature discusses its appropriateness for various problems and Shore and Johnson (1980) provide
a thorough axiomatic justification of its correctness. Other appearances of similar methods in econometrics include Golan, Judge and Miller (1996) and the Bayesian Method of Moments (BMOM) of Zellner (1997).

11 In fact $f^{(5)}(Z_{t-1}, Z_t)$ can be written as a Taylor series around $Z_{t-1} = Z_t = 1$ and the coefficient of $Z_{t-1}Z_t$ examined to give an approximate sense of the covariance implications of various parameter combinations.

12 Since $f^{(5)}$ is highly non-linear these pairs were solved by a numerical root-finding algorithm using a MINPACK subroutine via Gnu Octave.

13 For numerical convenience we have computed distributions which are truncated such that the support is the interval $(0.7, 1.3)$ which is equivalent to at least 5 standard deviations either side of the sample mean; for many practical purposes, therefore, this is a very close approximation to an untruncated distribution.

14 For actual computation we have used a slightly modified version of the algorithm proposed by Ormoneit and White (1999). Various subroutines from LAPACK and the Gnu Scientific Library (GSL) were used in C++ code.

15 Wolfowitz (1953) presented a theory of parameter estimation based on such a minimum-distance method, in which parameters should be adjusted until an optimal $p$-value is obtained and we note that it would be straightforward to carry-out such an optimisation process here. Although unusual in econometrics we have found several recent examples of similar approaches in the broader statistical literature: Kamungo and Zheng (2004) (noisy pattern recognition), Voss et al. (2004) (cognitive science) and Weber (2006). Consistency properties of such estimators are considered by Gyorfi et al. (1996).

16 Caution is required when interpreting KS statistics involving dependent data since the test assumes independence. Chicheportiche and Bouchaud (2011) analyze this problem and demonstrate how correct $p$-values can be computed by Monte-Carlo methods. We have not followed their specific approach here, but the results which they present lead us to believe that our $p$-values would be increased if we were to do so.

**Funding.** Thorp acknowledges Australian Research Council Discovery Grant 0877219. The Chair of Finance and Superannuation, UTS, receives support from the Sydney Financial Forum (through Colonial First State Global Asset Management), the NSW Government, the Association of Superannuation Funds of Australia (ASFA), the Industry Superannuation Network (ISN), and the Paul Woolley Centre for the Study of Capital Market Dysfunctionality, UTS.

**Acknowledgements.** Earlier versions of this work were presented at the Econometric Society Australasian Meeting, Brisbane, July, 2007 and the Symposium on Endowment Management, European Finance Association Meetings, Athens, August 2008. Comments of participants in these occasions are gratefully acknowledged.
References


Data Appendix

Wellcome Trust portfolio proxy returns data are monthly from January 1990 to July 2010. Total portfolio return is the log change of the weighted sum of monthly returns to each index and the cash rate less the log change in the inflation rate: \( \dot{p} = 0.5 \left[ \ln \left( \frac{\text{CPI}_t}{\text{CPI}_{t-1}} \right) + \ln \left( \frac{\text{Earnings}_t}{\text{Earnings}_{t-1}} \right) \right] \). All series are from DataStream apart from the Cash rate which is from the Bank of England database.

Table 3: Portfolio weights and sources

<table>
<thead>
<tr>
<th>Asset Class</th>
<th>Data</th>
<th>Mnemonic/Source</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 UK Equity</td>
<td>FTSE All Share</td>
<td>FTALLSH(RI)</td>
<td>32.2%</td>
</tr>
<tr>
<td>2 Global Equity</td>
<td>MSCI World ex UK</td>
<td>MSWFUK$(RI)^\sim$U$\sim$BPN using BBGBPSP(ER)</td>
<td>32.0%</td>
</tr>
<tr>
<td>3 Overseas Equity</td>
<td>MSCI Emerging Markets</td>
<td>MSEMKF$(RI)^\sim$U$\sim$BPN using BBGBPSP(ER)</td>
<td>5.0%</td>
</tr>
<tr>
<td>4 UK Gilts</td>
<td>BOFA</td>
<td>MLUK10(RI)$^\sim$£</td>
<td>2.8%</td>
</tr>
<tr>
<td>5 Property</td>
<td>IPD</td>
<td>UKIPDRI,F</td>
<td>7.5%</td>
</tr>
<tr>
<td>6 Hedge Funds</td>
<td>CSFB/Tremont Hedge Fund</td>
<td>CSTHEDG$^\sim$£</td>
<td>3.6%</td>
</tr>
<tr>
<td>7 Private Equity</td>
<td>UK Trusts Priv. Equity</td>
<td>ITVCAPT(RI)$^\sim$£</td>
<td>11.5%</td>
</tr>
<tr>
<td>8 Cash</td>
<td>3-month CD rate</td>
<td>Bank of England</td>
<td>5.4%</td>
</tr>
<tr>
<td>9 Inflation</td>
<td>Average of CPI and Earnings</td>
<td>CPI: UKCPHARMF</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wages: UKWAGES.E</td>
<td></td>
</tr>
</tbody>
</table>

Hedge fund data are the log of monthly net-of-fee returns to the Hedge Fund Research (HFR) equally-weighted Composite Index (HFRI), deflated using the same method as the Wellcome Trust data, January 1990-July 2010. In December 2009, the HFR index was constructed from 2,481 single manager funds, of which 1,930 are classified as active funds and 551 as ‘graveyard’ funds (funds that have gone out of business between January 1994 and December 2009 but whose track record remains in the database).