Center for Quantitative Economics

Estimating Continuous-Time Income Models

Christian Schluter[†] and Mark Trede[‡]

18/2011

[†] Department of Economics, University of Southampton, UK

[‡] Department of Economics, University of Münster, Germany

Estimating Continuous-Time Income Models

Christian Schluter* | Mark Trede†
University of Southampton | Universität Münster

January 2011

Abstract: A fundamental component of inter-temporal consumption-saving and portfolio allocation models is a statistical model of the income process. While income processes are commonly unobservable income flows which evolve in continuous time, observable income data are usually discrete, having been aggregated over time. We consider continuous-time earning processes, specifically non-linearly transformed Ornstein-Uhlenbeck processes, and the associated integrated, i.e. time aggregated process. Both processes are characterized, and we show that time aggregation alters important statistical properties. The parameters of the earning process are estimable by GMM, and the finite sample properties of the estimator are investigated. Our methods are applied to annual earnings data for the US. It is demonstrated that the model replicates well important features of the earnings distribution.

Keywords: income processes, integrated non-linearly transformed Ornstein-Uhlenbeck process, temporal aggregation.

JEL classification: C22, E21, E24, J31

*Department of Economics, University of Southampton, Highfield, Southampton, SO17 1BJ, UK. Tel. +44 (0)2380 59 5909, Fax. +44 (0)2380 59 3858. Email: C.Schluter@soton.ac.uk. http://www.economics.soton.ac.uk/staff/schluter/

†Corresponding author. Center for Quantitative Economics and Center for Nonlinear Science, Westfälische Wilhelms-Universität Münster, Am Stadtgraben 9, 48143 Münster, Tel.: +49-251-83 25006, Fax: +49-251-83 22012, Email: mark.trede@uni-muenster.de.

1 Introduction

A fundamental component of inter-temporal consumption-saving and portfolio allocation models is a statistical model of the income process (see e.g. the discussion in Wang (2009)). Assumed to evolve in continuous time, popular modeling choices are variants of Brownian motion processes.[1] The principal obstacle to the empirical implementation and the eventual testing of a model's prediction, however, is the nature of the income data reported in the usual surveys: neither is the income flow observable, nor is the income process sampled at specific time points; what is reported in survey data is income data aggregated, by necessity, over time intervals. We address this problem by considering a statistical model of a continuous-time earnings process, and we propose methods to estimate its parameters from discretely sampled time-aggregated data. The estimation approach suggested in this paper therefore bridges the gap between the theoretical models and their empirical application. Moreover, we consider (Mincerian) classes of income processes, which are a more general than common modeling choices, and show that the estimated process describes well the US earnings distributions.

More specifically, we assume that the unobserved continuous-time earnings process is a non-linearly transformed Ornstein-Uhlenbeck (OU) process. In the baseline model, the transformation is an exponentiation, in the generalized model it is an inverted Box-Cox transformation. This process is sampled over possibly non-regular intervals, resulting for the baseline model in an integrated exponentiated Ornstein-Uhlenbeck process (intexpOU) and for the general model in an integrated inverted Box-Cox Ornstein-Uhlenbeck process (intinvBCOU). We characterize both expOU and intexpOU processes in terms of distributions and moments. In particular, we show that whereas the expOU process is Markov and lognormal, the intexpOU is neither. This is an important finding since it demonstrates that wrongly assuming the observed integrated process to have the same distributional properties as the unobservable underlying continuous-time process would introduce a temporal aggregation bias.

We demonstrate how the parameters of the unobservable income process are estimable from standard time-aggregated data by means of a GMM procedure. The merit of our approach is illustrated using US PSID income data, and we show that the estimated model fits the data very well.

This is the first paper, to the best of our knowledge, to consider the estimation of continuous-time earnings processes from time-aggregated data. The common approach in the labor eco-

[1] E.g. simple geometric Brownian motion (e.g. Bodie et al. (1992), Koo (1998), or Bick et al. (2009)), Brownian motion with drift (Henderson, 2005), or geometric Brownian motion with time-varying drift and depending on other economic variables (Munk and Sørensen, 2009). Wang (2004, 2006, 2009) considers an Ornstein-Uhlenbeck process.

nomics literature is to estimate a discrete-time error component model on annual earnings data (e.g. MaCurdy, 1982, Abowd and Card, 1989, Baker, 1997, Guvenen, 2009). Integrated diffusion processes are thus new to this empirical setting, but have been considered in other fields. We highlight the principal differences. In the statistics literature, Gloter (2001) considers an integrated stationary Ornstein-Uhlenbeck process (intOU), which is shown to be Gaussian and ARMA(1,1) with an exponentially decaying α-mixing coefficient. The likelihood is intractable and he proposes a Whittle estimator for the OU parameters. However, his results do not apply to our case since the non-linear transform (exponentiation in the baseline model) prior to integration leads to completely different distributional properties of intexpOU and intOU processes.[2] Such intOU processes are considered particularly in finance, and estimated for settings in which the sampling time interval, required to be regular, converges to zero (e.g. Ditlevsen and Sørensen (2004), or Gloter (2006)). The leading application is stochastic volatility modeling in finance (e.g. Barndorff-Nielsen and Sheppard (2001)). In contrast to this literature our empirical setting does not allow to shrink the time interval to zero, nor do we require the time intervals to be regular. Another strand of the finance literature deals with integrated continuous-time processes in the context of Asian option pricing (e.g. Carr and Schröder (2004)). Finally, turning to the consequences of temporal aggregation, such aggregation can, as in our case, lead to important differences between the continuous and the integrated process. Referring to such differences as time aggregation biases, these have been studied in a macro context in Harvey and Stock (1989) and Christiano, Eichenbaum and Marshall (1991), who consider how time-aggregation alters the results of tests of the permanent income hypothesis.

This paper is structured in the following way: In Section 2 we present the statistical model for the income process. Section 3 derives the moments of the time-aggregated, observable process. In Section 4 we suggest a more general non-linear transformation (Box Cox transformation) for the income flow which permits the modeling of heavy tails. Section 5 sets out the GMM estimation procedure. While the asymptotic properties of GMM are known to be attractive, not much can be said about its small sample properties. Therefore, in Section 6, we conduct simulation exercises and find that our estimation approach performs well even in relatively small samples. Section 7 contains the empirical application in which we estimate the parameters of the continuous-time model using annual panel data from the US PSID. The model turns out to fit the data very well. Section 8 concludes. All proofs are collected in the Appendix.

[2] Other related work includes Bhattacharya, Thomann and Waymire (2001), who derive partial differential equations for the distribution of integrals of geometric Brownian motions. This is a special case of the intexpOU process if the mean reversion parameter of the OU process is absent. Comte, Genon-Catalot and Rozenholc (2009) suggest a nonparametric estimation method for integrated diffusions.

2 The Statistical Model for the Income Process

The log-income flow of an individual, denoted by $\{\ln Y(t) : t \geq t_0\}$, is assumed to follow a stochastic process evolving in continuous time. t_0 denotes the starting time of the process. We impose a structure on this earnings process which follows conventional modeling, except that our process evolves in continuous rather than discrete time. Specifically, we assume that the earnings process decomposes additively into two independent parts. The error process is Gaussian and denoted by $\{u(t) : t \geq t_0\}$, the model for the mean log-income flow is denoted by $\{\tilde{y}(t) : t \geq t_0\}$, and we assume that

$$\ln Y(t) = \tilde{y}(t) + u(t). \tag{1}$$

In Section 4 below we also consider more general Box-Cox transformations of the $Y(t)$ process.

The model for $\{\tilde{y}(t)\}$ is standard in the sense of relating the mean log-income linearly to observables such as measures of human capital. In our empirical application we postulate a Mincerian model. For notational convenience later, we partition the relevant observables into time-invariant and time-varying covariates

$$\tilde{y}(t) = m + Z_1^\top \beta + Z_2^\top(t)\gamma. \tag{2}$$

We treat the regressors as exogenous and follow e.g. Abowd and Card (1989) in ignoring the potential endogeneity of the human capital measure such as schooling which could arise from unobserved ability. The intercept m, allowed to be individual-specific in order to accommodate unobserved heterogeneity, is modelled as a random effect,

$$m = \mu + \varepsilon,$$

with ε having a zero-mean Gaussian distribution with variance σ_ε^2.

The Gaussian latent variable or error process $\{u(t)\}$ is assumed to be a zero-mean Ornstein-Uhlenbeck (OU) process, governed by the stochastic differential equation

$$du(t) = -\eta u(t)\,dt + \sigma dW(t), \tag{3}$$

with solution

$$u(t) = u(t_0)e^{-\eta t} + \sigma \int_{t_0}^{t} e^{\eta(s-t)} dW(s). \tag{4}$$

$\{W(t) : t \geq t_0\}$ is the standard Wiener process, and $\eta \in \mathbb{R}$ and $\sigma > 0$ are the parameters of the process.

The OU process is an attractive point of departure for two reasons. First, it is the continuous-time counterpart of an autoregressive process in discrete time. Autoregressive processes are commonly used for models of income dynamics in discrete time. Second, OU processes capture not only stable processes but also unit root processes and, if $\eta < 0$, even explosive processes. In contrast to the common stability assumption, we do not impose any restrictions on the parameter η.

The parameter σ determines the strength of the stochastic income component, and equals the diffusion coefficient of the process. $-\eta u(t)$ is the instantaneous mean of the OU process. The start value of the OU process $u(t_0)$ is assumed to be stochastic with $E(u(t_0)) = 0$ and $Var(u(t_0)) = s_0^2$. Mean and covariances are given by $E\{u(t)\} = 0$ and, for $\eta \neq 0$,[3]

$$\sigma_{s,t} \equiv Cov(u(s), u(t)) = \frac{\sigma^2}{2\eta} e^{-\eta|t-s|} + ae^{-\eta(t+s)} \tag{5}$$

with

$$a = s_0^2 - \frac{\sigma^2}{2\eta} \tag{6}$$

for notational simplicity. Finally we note that the OU process is weakly stationary if $a = 0$, i.e. $s_0^2 = \sigma^2\eta^{-1}/2$, and strongly stationary if $u(t_0)$ is in addition Gaussian. However, we do not impose any stationarity assumption.

Exponentiating the process (1) yields the continuous-time income process $\{Y(t) : t \geq t_0\}$,

$$Y(t) = \exp(\tilde{y}(t) + u(t)),$$

which is an exponentiated OU process (expOU).

A key assumption is that the econometrician cannot sample the process at specific points of time. Instead, only the time-aggregated process is observable, i.e. the econometrician observes the integrated process for non-overlapping time intervals $[t_0, t_1], \dots, [t_{T-1}, t_T]$. Depending on the specific application, these time intervals could be individual-specific if the data are spell data, and common across individuals when the data are annual panel data. Below we refer to time intervals that are common across individuals and of the same length Δ as regular intervals. The observable process is thus the integrated exponential Ornstein-Uhlenbeck (intexpOU) process

$$S_k = \int_{t_{k-1}}^{t_k} Y(t)\, dt = \int_{t_{k-1}}^{t_k} e^{\tilde{y}(t)} e^{u(t)} dt \qquad k = 1, \dots, T. \tag{7}$$

For ease of reference, we collect the model parameters in the vector

$$\theta = [\eta, \sigma, s_0, \mu, \sigma_\varepsilon, \beta, \gamma]^\top. \tag{8}$$

We proceed to characterize both the unobservable income process $\{Y(t)\}$ and the observable integrated process $\{S_k\}$.[4] In particular, important distributional properties of $\{Y(t)\}$ will not be inherited by $\{S_k\}$.

[3] We exclude the unit root case for expositional and notation ease by assuming $\eta \neq 0$. However, all results can be specialized for $\eta \to 0$.

[4] We follow notational convention and denote a continuous-time stochastic process by $A(t)$ and an aggregated or discrete-time process by B_k.

3 Characterizing the Time Aggregated Process

Proposition 1 *The unobservable income process $\{Y(t)\}$ is lognormal and Markov, with*

$$E(Y(t)) = \exp\left(\mu + Z_1^\top \beta + \frac{\sigma_\varepsilon^2}{2} + \frac{\sigma^2}{4\eta}\right) \exp\left(Z_2^\top(t)\gamma + a\frac{e^{-2\eta t}}{2}\right), \tag{9}$$

$$\begin{aligned} E(Y(s)Y(t)) &= \exp\left(2\left(\mu + Z_1^\top \beta + \sigma_\varepsilon^2\right)\right) \times \\ &\exp\left(\left[Z_2^\top(s) + Z_2^\top(t)\right]\gamma\right) \times \\ &\exp\left(\frac{\sigma^2}{2\eta}\left[1 + e^{-\eta|t-s|}\right] + ae^{-\eta(t+s)} + \frac{a}{2}\left(e^{-2\eta s} + e^{-2\eta t}\right)\right). \end{aligned} \tag{10}$$

The integrated process does not inherit these distributional properties:

Lemma 2 *The observable process $\{S_k : k = 1, \ldots, T\}$ is neither Markov nor lognormal.*

We therefore consider the moments of the aggregated process, which follow from an application of Fubini's theorem.

Corollary 3 *The moments of the observable process $\{S_k : k = 1, \ldots, T\}$ are given by*

$$\begin{aligned} E(S_k^n) &= E\left(\left(\int_{t_{k-1}}^{t_k} Y(t)\,dt\right)^n\right) \\ &= \exp\left(n\left(\mu + Z_1^\top \beta\right) + \frac{1}{2}n^2\sigma_\varepsilon^2\right) \times \\ &\int_{t_{k-1}}^{t_k} \cdots \int_{t_{k-1}}^{t_k} \exp\left(\sum Z_2^\top(s_i)\gamma\right) \\ &\times \exp\left(\frac{1}{2}\sum \sigma_{s_i,s_i} + \sum\sum_{j>i} \sigma_{s_i,s_j}\right) ds_1 \ldots ds_n \end{aligned} \tag{11}$$

where $\sigma_{s,t}$ is given by (5) and n is an integer.

Lemma 4 *Moments of S_k of all orders exists, and the distribution of S_k therefore cannot be heavy-tailed.*

Recall that a heavy-tailed distribution is one whose tail decays like a power function, i.e. $1 - F(x) = x^{-1/\gamma}L_0(x)$ for sufficiently large x, where L_0 is a slowly varying function and $\gamma > 0$. In Section 4, we therefore consider a generalized model which subsumes the intexpOU and heavy-tailed processes as special cases.

Corollary 5 *The mixed moments for intervals k and r are*

$$E(S_k S_r) = \int_{t_{k-1}}^{t_k} \int_{t_{r-1}}^{t_r} E(Y(s)Y(t))\,dsdt \tag{12}$$

with integrand given by (10).

Without giving more structure to the covariate process, we cannot characterize the moments any further. A special case arises when the covariate process is absent. We then can state exact expressions for the first moment, and approximations for the covariances, which give some insights into the behavior of the more general process. For expositional brevity we focus on regular intervals of length Δ. In Appendix A we derive the following statements.[5] The first moment satisfies exactly

$$E\left(S_k\right) \times \exp\left(-\frac{\sigma^2}{4\eta}\right) = \Delta + \frac{1}{2\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{a}{2}e^{-2\eta(k-1)\Delta}\right]^i\left[1-e^{-2i\eta\Delta}\right],$$

and, for $\eta > 0$,[6] we have

$$\begin{aligned}
& Cov\left(S_l, S_k\right)_{l<k} \times \exp\left(-\frac{\sigma^2}{2\eta}\right) \\
\simeq\; & \frac{1}{\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{\sigma^2}{2\eta}\right]^i\frac{1}{i\eta}\left[e^{i\eta\Delta}-1\right]\left[1-e^{-i\eta\Delta}\right]e^{-i\eta(k-l)\Delta} \\
& -\frac{1}{2\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{a}{2}\right]^i\frac{1}{2i\eta}e^{-2i\eta(l-1)\Delta}\left[e^{-2i\eta\Delta}-1\right]^2e^{-2i\eta(k-1)\Delta},
\end{aligned}$$

and

$$\begin{aligned}
& Var\left\{S_k\right\} \times \exp\left(-\frac{\sigma^2}{2\eta}\right) \\
\simeq\; & \frac{2}{\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{\sigma^2}{2\eta}\right]^i\left(\Delta+\frac{1}{i\eta}\left[e^{-i\eta\Delta}-1\right]\right) \\
& -\frac{1}{2\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{a}{2}\right]^i\frac{1}{2i\eta}\left[e^{-2i\eta\Delta}-1\right]^2e^{-4i\eta(k-1)\Delta}.
\end{aligned}$$

These two approximations become exact in the stationary case with $a = 0$. The expressions highlight the effect of temporal aggregation. In particular mean and variances become unbounded as $\Delta \to \infty$. The covariances shrink to zero for fixed l as $k \to \infty$. Covariances and variances approximations increase in the diffusion coefficient σ^2 for $a \geq 0$, while the effect of increasing η is ambiguous. For $a > 0$ the covariance approximations are decreasing in s_0 while the first moments are increasing.

3.1 Digression: The intexpOU Process and standard Error Component Modeling

We consider the relationship between the structural equation (7) describing the intexpOU process and the estimating equations of the error component modeling (ECM) approach as commonly

[5] In our numerical verifications, using the first term of the summation yields already good results. Note, however, that in the simulation study below we use the exact expressions.

[6] Similar expressions can be derived for $\eta < 0$.

implemented in the literature. Recall that this consists in first filtering out observables using a linear regression of log-income on observables, and then to estimate an error component model using the empirical covariance structure of the residuals in the second step.

For expositional simplicity assume that $Z_2(t)$ only contains an aggregate time effect and an age effect, so $Z_2^\top(t)\gamma = \gamma_1 t + \left[age_{t_{k-1}} + t - t_{k-1}\right]\gamma_2$, and that the time intervals are regular and of length Δ (usually a 'year'). We have

$$\begin{aligned}\log(S_k) &= \mu + Z_1^\top\beta + \gamma_1 t_{k-1} + \gamma_2 age_{t_{k-1}} + \varepsilon \\ &\quad + \log\int_0^\Delta \exp\left((\gamma_1+\gamma_2)\tau + u(t_{k-1}+\tau)\right)d\tau.\end{aligned}$$

Consider the expectations of the last term.

$$\mu_k \equiv E\left\{\log\int_0^\Delta \exp\left((\gamma_1+\gamma_2)\tau + u(t_{k-1}+\tau)\right)d\tau\right\}.$$

Using the strict concavity of the log function we observe that

$$\begin{aligned}&\log\int_0^\Delta \exp\left((\gamma_1+\gamma_2)\tau + u(t_{k-1}+\tau)\right)d\tau \\ > &\int_0^\Delta \log\exp\left((\gamma_1+\gamma_2)\tau + u(t_{k-1}+\tau)\right)d\tau \\ = &\frac{1}{2}(\gamma_1+\gamma_2)\left[\Delta_k^2\right] + \int_{t_{k-1}}^{t_k} u(t)\,dt,\end{aligned}$$

which implies that $\mu_k > 0$ since $u(t)$ has mean zero. The estimating equation used by the standard ECM approach is therefore

$$\begin{aligned}\log(S_k) &= (\mu+\mu_k) + Z_1^\top\beta + \gamma_1 t_{k-1} + \gamma_2 age_{t_{k-1}} + res_k \\ res_k &= \varepsilon + \left[\log\int_0^\Delta \exp\left((\gamma_1+\gamma_2)\tau + u(t_{k-1}+\tau)\right)d\tau - \mu_k\right]\end{aligned} \tag{13}$$

where the true residual res_k has mean zero.

Lemma 6 *The marginal effects of the time-invariant covariates, age, and the time effect in the first-step ECM regression equal the true population coefficients β and γ_2.*

Next, consider the covariance structure of the residuals res_k, upon which the second step of the ECM approach consists in imposing a model. A common choice is to model the residual as the sum of a random effect ε, a random walk part p_k representing a permanent shock, and a MA(1) component z_k representing temporary shocks: $res_{k,ECM} = \varepsilon + p_k + z_k$ where ε also appears in the reduced form (13), with $p_k = p_{k-1} + w_k$ and $z_k = x_k - \delta x_{k-1}$ and $w_k \sim^{iid} \left(0, \sigma_w^2\right)$ and $x_k \sim^{iid} \left(0, \sigma_x^2\right)$. Upon taking first difference using the ECM we have

$$\Delta r_{k,ECM} \equiv r_k - r_{k-1} = w_k + x_k - (1+\delta)\,x_{k-1} + \delta x_{k-2}$$

so that

$$E\left\{\Delta r_{k,ECM}\Delta r_{k+s,ECM}\right\} = \begin{cases} \sigma_w^2 + \left[1 + (1+\delta)^2 + \delta^2\right]\sigma_x^2 & \text{for } s = 0 \\ -(1+\delta)^2\sigma_x^2 & \text{for } s = 1 \\ \delta\sigma_x^2 & \text{for } s = 2 \\ 0 & \text{for } s > 2. \end{cases} \tag{14}$$

However, using (13) we have

$$\Delta r_k = \log\int_0^{\Delta} \exp\left(\left(\gamma_1 + \gamma_2\right)\tau + u_i\left(t_{k-1} + \tau\right)\right)d\tau - \mu_k,$$

so $E\left\{\Delta r_k \Delta r_{k+s}\right\}$ is a complicated function of the structural parameters $\left(\eta, \sigma^2, \gamma\right)$, and is not available in closed form. In general, the ECM estimating equations (14) do not describe correctly the structure $E\left\{\Delta r_k \Delta r_{k+s}\right\}$.

4 A Generalized Box-Cox Transformed Model

The Mincerian formulation of the income process (1) could be criticized for two reasons. First, the logarithmic transformation is a fairly ad hoc assumption. Second, as Lemma 4 states, the distribution of time aggregated income S_k cannot exhibit heavy tails since all moments exist. Yet it is an enduring stylized fact going back to Pareto (1896) that some income and earnings distributions exhibit heavy tails, i.e. the tails of the distribution decay like power functions (see e.g. Schluter and Trede 2002, 2008). Common earnings models in the literature fail to generate these, and so does the cross-sectional earnings distribution implied by the intexpOU process. It is therefore desirable to seek a generalization of the current model which optimally determines the transformation of $Y(t)$ in a data-dependent manner, and which nests the Mincerian logarithmic transformation and heavy tail transformations as special cases. These desiderata are fulfilled by the Box Cox transformation g_λ given by

$$g_\lambda(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \ln(x) & \text{otherwise} \end{cases},$$

for $x > 0$ leading to the generalized income model[7]

$$Y(t) = g_\lambda^{-1}\left(\tilde{y}(t) + u(t)\right).$$

In our application the Box Cox parameter λ will be estimated. The next Lemma elucidates the role played by λ.

[7] Note that since the argument of the Box Cox transformation is required to be positve, this implies that the distribution of $u(t)$ be suitably truncated when $\lambda < 0$ in any formal analysis.

Lemma 7 *The model is linear for $\lambda = 1$, the Mincerian model follows for $\lambda = 0$. For $\lambda < 0$ the distribution of $Y(t)$ exhibits a heavy tail, whereas for $\lambda \geq 0$ the right tail is decaying exponentially fast. Expressed equivalently $F_{Y(t)}$ is regularly varying in the right tail of its support when $\lambda < 0$, and slowly varying otherwise. When $\lambda < 0$, the tail index (or the reciprocal of the coefficient of regular variation) is proportional to $|\lambda|^{-1}$.*

The generalized observable integrated process is now given by

$$S_k^{(\lambda)} = \int_{t_{k-1}}^{t_k} Y(t)\,dt = \int_{t_{k-1}}^{t_k} g_\lambda^{-1}\left(\tilde{y}(t) + u(t)\right) dt$$

which we refer to below as the intinvBCOU process. The intexpOU process is $S_k = S_k^{(0)}$.

Proposition 1 extends naturally to this setting, for instance the first moment being $E(S_k^{(\lambda)}) = \int_{t_{k-1}}^{t_k} E\left(g_\lambda^{-1}\left(\tilde{y}(t) + u(t)\right)\right) dt$. The non-separability of $g_\lambda^{-1}\left(\tilde{y}(t) + u(t)\right)$ can be extremely costly in terms of computation time. For many practical applications it is therefore advisable to use higher order Taylor series approximations. Let $c(t) = \lambda\left[\mu + Z_1^\top \beta + Z_2^\top(t)\,\gamma\right] + 1$, then for $\lambda \neq 0$ we have, correct to fourth order,

$$\begin{aligned}
& E(Y(t)) \\
= \; & c(t)^{1/\lambda} + \frac{1}{2}(1-\lambda)\,c(t)^{1/\lambda-2}\left(\sigma_\varepsilon^2 + \sigma_{t,t}\right) \\
& + \frac{1}{24}(1-\lambda)(1-2\lambda)(1-3\lambda)\,c(t)^{1/\lambda-4}\left(3\sigma_\varepsilon^4 + \sigma_{t,t}^2 + 6\sigma_\varepsilon^2\sigma_{t,t}\right)
\end{aligned}$$

and

$$\begin{aligned}
& E(Y(s)\,Y(t)) \\
= \; & [c(s)\,c(t)]^{1/\lambda} + [c(s)\,c(t)]^{1/\lambda-1}\left(\sigma_\varepsilon^2 + \sigma_{s,t}\right) \\
& + \frac{1-\lambda}{2}\left(\sigma_\varepsilon^2 + \sigma_{t,t}\right) c(t)^{1/\lambda-2}\,c(s)^{1/\lambda} + \frac{1-\lambda}{2}\left(\sigma_\varepsilon^2 + \sigma_{s,s}\right) c(s)^{1/\lambda-2}\,c(t)^{1/\lambda} \\
& + \left[\frac{1-\lambda}{2}\right]^2 [c(s)\,c(t)]^{1/\lambda-2}\left(3\sigma_\varepsilon^4 + \sigma_\varepsilon^2\left(\sigma_{t,t} + \sigma_{s,s} + 4\sigma_{s,t}\right) + 2\sigma_{s,t}^2 + \sigma_{s,s}\sigma_{t,t}\right) \\
& + \frac{1-\lambda}{2}(1-2\lambda)\,c(t)^{1/\lambda-3}\,c(s)^{1/\lambda-1}\left(\sigma_\varepsilon^4 + \sigma_\varepsilon^2\left(\sigma_{t,t} + \sigma_{s,t}\right) + \sigma_{s,t}\sigma_{t,t}\right) \\
& + \frac{1-\lambda}{2}(1-2\lambda)\,c(s)^{1/\lambda-3}\,c(t)^{1/\lambda-1}\left(\sigma_\varepsilon^4 + \sigma_\varepsilon^2\left(\sigma_{s,s} + \sigma_{s,t}\right) + \sigma_{s,t}\sigma_{s,s}\right) \\
& + \frac{1-\lambda}{24}(1-2\lambda)(1-3\lambda)\,c(t)^{1/\lambda-4}\,c(s)^{1/\lambda}\left(3\left(\sigma_\varepsilon^4 + \sigma_{t,t}^2\right) + 6\sigma_\varepsilon^2\sigma_{t,t}\right) \\
& + \frac{1-\lambda}{24}(1-2\lambda)(1-3\lambda)\,c(s)^{1/\lambda-4}\,c(t)^{1/\lambda}\left(3\left(\sigma_\varepsilon^4 + \sigma_{s,s}^2\right) + 6\sigma_\varepsilon^2\sigma_{s,s}\right)
\end{aligned}$$

where $\sigma_{s,t}$ are given by (5).

5 Estimation and Inference

Inspection of equations (11) and (12), and hence of (9) and (10), makes it clear that all parameters $\theta = \left[\eta, \sigma, s_0, \mu, \sigma_\varepsilon^2, \beta, \gamma\right]^\top$ are identified provided that first and second moments are employed in the estimation. We estimate the parameters in particular by iterated GMM using estimating functions based on the first and mixed moments. In addition, orthogonality conditions for the covariates are included.

More specifically, denote by $S_{k,i}$ the observed intexpOU process for individual $i = 1, \ldots, N$ for the sampling interval $k = 1, \ldots, K$. Similarly, define $Z_{1,i}$ as the time-invariant covariates (including a constant 1 for the intercept) and $Z_{2,i}(t)$ as the time-varying covariates of individual i. The first moment conditions for sampling interval k are

$$f_{1,k} = \frac{1}{N}\sum_{i=1}^{N} Z_{k,i}S_{k,i} - \frac{1}{N}\sum_{i=1}^{N} Z_{k,i}E_\theta\left(S_{k,i}\right) = 0 \tag{15}$$

where $Z_{k,i}^\top = [Z_{1,i}^\top, Z_{2,i}^\top(t_{k-1})]$ is the vector of covariates; the time-varying $Z_{2,i}$ is approximated by its value at the start of the interval. The dimension of $f_{1,k}$ depends on the number of covariates; if covariates are absent, $f_{1,k}$ is simply the difference between the empirical and theoretical first moment.

In order to identify the parameters of the diffusion process $\{u(t)\}$, we need a set of second moment conditions. We use

$$\begin{aligned} f_{2,k} &= \left(\frac{1}{N}\sum_{i=1}^{N} S_{1,i}S_{k,i} - \bar{S}_1\bar{S}_k\right) - \left(\frac{1}{N}\sum_{i=1}^{N} E_\theta\left(S_{1,i}\right)E_\theta\left(S_{k,i}\right) - \overline{E_\theta S_1}\cdot\overline{E_\theta S_k}\right) \\ &= 0 \end{aligned}$$

where

$$\bar{S}_k = \frac{1}{N}\sum_{i=1}^{N} S_{k,i} \quad \text{and} \quad \overline{E_\theta S_k} = \frac{1}{N}\sum_{i=1}^{N} E_\theta\left(S_{k,i}\right)$$

for $k = 1, \ldots, K$. Hence, $f_{2,k}$ is the difference between the empirical and the theoretical autocovariance of order $k-1$.

Stack $f_{1,k}$ for all the different sampling intervals to get $f_1^\top = \left[f_{1,1}^\top, \ldots, f_{1,K}^\top\right]$, a vector of length $K \cdot P$ where P is the number of covariates. Stack similarly the estimating function for the second moments to get $f_2^\top = [f_{2,1}, \ldots, f_{2,K}]$, and finally stack these to get $f^\top = \left[f_1^\top, f_2^\top\right]$, which is a vector of length $K(P+1)$. Where necessary we make the dependence on θ explicit by writing $f(\theta)$. Denote by $\Omega(\theta)$ the theoretical covariance matrix of $f(\theta)$.

The initial or unweighted estimate of θ is obtained by unweighted GMM, $\theta_1 = \arg\min f(\theta)^\top f(\theta)$. The n^{th} iterate is obtained as

$$\theta_n = \arg\min f(\theta)^\top \Omega\left(\theta_{n-1}\right)^{-1} f(\theta),$$

and we iterate until the parameter vector has converged. Denote the converged value by $\widehat{\theta}$.

More generally, write the estimate of the criterion function as $\widehat{Q}(\theta) = f(\theta)^{\top} \widehat{W} f(\theta)$ where $\widehat{W}$ be a weighting matrix which converges in probability to a positive semi-definite matrix W. The estimator is given by $\widehat{\theta} = \arg\min f(\theta)^{\top} \widehat{W} f(\theta)$. Under standard regularity conditions (e.g. Theorem 3.2 of Newey and McFadden, 1994) the estimator satisfies

$$\left(\widehat{\theta} - \theta_{\text{true}}\right) \xrightarrow{d} N\left(0, \left(G^{\top} W G\right)^{-1} G^{\top} W \Omega W G \left(G^{\top} W G\right)^{-1}\right)$$

where θ_{true} is the population value, $G(\theta) = df(\theta)/d\theta$, $G = G(\theta_{\text{true}})$, $\Omega(\theta)$ is the theoretical covariance matrix of $f(\theta)$ and $\Omega = \Omega(\theta_{\text{true}})$. Let $\widehat{\Omega}$ denote its estimator. The efficient weighting matrix is Ω^{-1}, estimated in the n^{th} iteration by $\Omega(\widehat{\theta}_{n-1})^{-1}$.

5.1 Initializations

The structural model permits convenient initializations of parameters β and σ_{ε}^2. In particular the coefficients for the time-invariant covariates β can be estimated by a simple regression if the time-varying covariates have a simple structure. As in our empirical application, assume that the time-varying covariates $Z_{2,i}(t)$ consist of a time effect and a polynomial in age. Then we can consider groups defined by ages, difference out group invariants, and identify the coefficients of Z_1 by with-in group variations in Z_1.

More specifically, for interval k, the first moment can be written as

$$E(S_{i,k}) = \exp\left(\mu + Z_{1,i}\beta + \sigma_{\varepsilon}^2/2 + \sigma^2/4\eta\right) \Psi_{age(i),k}$$

with $\Psi_{age(i),k} = \int_{t_{k-1}}^{t_k} \exp(Z_{2,i,t}(t)\gamma)\, dt$ being the same for individuals of the same age. Since $S_{i,k} = E(S_{i,k}) + \widetilde{error}_{i,k}$, a first order Taylor expansion yields $\log(S_{i,k}) \approx \mu + Z_{1,i}\beta + \sigma_{\varepsilon}^2/2 + \sigma^2/4\eta + \log \Psi_{1,age(i),k} + error_{i,k}$. For each age group (and each k), substract group means, and finally regress individual within-group deviation from group means of $\log(S_{i,k})$ on within-group deviation from group means of $Z_{1,i}$ to obtain β.

A good initial simple estimate of σ_{ε}^2 can be obtained in situations in which $\sigma_{s,t} \approx 0$, which requires $\eta > 0$ and $t \gg s$. For instance, in the empirical application first and last periods (indicated by l) are 7 years apart. In these circumstances group individuals present in periods 1 and l into cells defined by unique values of Z_1 and birth years. Then using (9) and (10), for each such cell c and $i \in c$ we have $E(S_{1,i}S_{l,i}) / [E(S_{1,i}) E(S_{l,i})] = \exp\left(\sigma_{\varepsilon}^2\right)$, and averaging over all cells yields an initial estimate of $\exp\left(\sigma_{\varepsilon}^2\right)$.

6 Simulation Evidence

We briefly investigate the finite sample performance of the GMM procedure in two different settings characterized by contrasting parameter values and lengths of time aggregation. Since the parameters $\theta = [\eta, \sigma, s_0]^\top$ of the diffusion process $\{u(t)\}$ are the principal objects of interest, we assume that the process $\{\tilde{y}(t)\}$ is absent. Throughout the experiments, time evolves on the unit time interval, and we assume that all individuals have the same start time $t_0 = 0$. The econometrician only observes the intexpOU processes $S_{k,i}$ for individuals $i = 1, \ldots, N$ and time intervals $k = 1, \ldots, K$. We consider initially $K = 5$ time intervals of varying lengths. In particular, let I_1 denote the intervals given by $[.2, .3]$, $[.3, .4]$, $[.4, .5]$, $[.5, .7]$, and $[.7, .9]$. Hence this setting illustrates the virtue of the model to accommodate sampling intervals of varying length, and substantial time aggregation. I_2 collects the time intervals $[.1, .15]$, $[.15, .2]$, $[.2, .25]$, $[.25, .3]$, $[.3, .35]$. Compared to the previous setting, the intervals are regular and time aggregations less substantial. The total number of first and all mixed moments given by (11) and (12) is thus 20, and we use all of these. The experiments have been repeated 500 times, we consider samples of sizes $N = 100$ and $N = 500$ (far smaller than the sample size of our empirical application), and compare results for iteration steps 0 (the unweighted estimation), 1 and 3. In the first experiment, we set $\theta = [2.3, .707, .2828]^\top$ – which happens to coincide with the estimates in our empirical application – and we report the estimates for both I_1 and I_2. In the next experiments, we consider contrasting parameter configurations and interval settings. In particular, in the second experiment we let $\theta = [1.5, .8, .5]^\top$ and estimate the model on I_1, in the third experiment we estimate $\theta = [.2, .2236, .3536]^\top$ on I_2. Note that in the second experiment all parameters are large compared to the third experiment, and that $a = s_0^2 - \sigma^2/(2\eta) > 0$ so the model is not stationary (but still stable), whereas $a = 0$ in the population in the third experiment (but we do not impose this restriction in the estimation).

Tables 1 and 2 reports the means and standard deviations (SD) of the parameter estimates across the simulations. We report these for the unweighted initialization step 0, and iterations $i = 1$ and $i = 3$ in which the weighting matrix is $\widehat{\Omega}(\widehat{\theta}_{i-1})^{-1}$.[8] $SE(\widehat{\theta}_3)$ is the mean standard error (SE) in iteration 3. If convergence is achieved in step 2 then this result is reported. In Table 1 we report the results for sample sizes 100 and 500, but all subsequent tables report results for $N = 500$ for reasons of space.

We turn to the results. Across all experiments we observe the following general features:

[8]Due to inaccuracies in the numerical approximation of the integrals in (11) and (12), the inverse of $\widehat{\Omega}(\widehat{\theta}_{i-1})$ may have some slightly negative eigenvalues, i.e. it is not always positive definite. In these circumstances, we render the weighting matrix positive definite by applying a spectral decomposition in which all eigenvalues below a tolerance value of 10^{-6} are equated to this value.

		$[\eta, \sigma, s_0] = [2.3, .707, .2828]$							
		I_1				I_2			
		iteration				iteration			
N		0	1	3	$SE(\widehat{\theta}_3)$	0	1	3	$SE(\widehat{\theta}_3)$
500	$\widehat{\eta}$	2.118	2.326	2.310	0.160	2.074	2.312	2.307	0.191
	(SD)	(0.989)	(0.355)	(0.228)	(0.007)	(0.762)	(0.292)	(0.224)	(0.008)
	$\widehat{\sigma}$	0.673	0.705	0.708	0.019	0.657	0.702	0.705	0.014
	(SD)	(0.207)	(0.033)	(0.031)	(0.002)	(0.170)	(0.021)	(0.018)	(0.000)
	$\widehat{s}_0$	0.253	0.273	0.278	0.035	0.279	0.283	0.285	0.019
	(SD)	(0.144)	(0.060)	(0.052)	(0.006)	(0.085)	(0.033)	(0.026)	(0.001)
100	$\widehat{\eta}$	2.133	2.600	2.362	0.361	2.310	2.576	2.319	0.431
	(SD)	(2.617)	(1.872)	(0.559)	(0.071)	(2.826)	(2.487)	(0.578)	(0.091)
	$\widehat{\sigma}$	0.634	0.698	0.705	0.045	0.628	0.684	0.702	0.031
	(SD)	(0.449)	(0.115)	(0.069)	(0.010)	(0.453)	(0.081)	(0.045)	(0.003)
	$\widehat{s}_0$	0.304	0.272	0.281	0.113	0.280	0.299	0.286	0.044
	(SD)	(0.584)	(0.179)	(0.126)	(0.130)	(0.399)	(0.181)	(0.060)	(0.017)

Table 1: Simulation evidence I: Estimating the parameters of the intexpOU process.

The estimation procedure copes well across different sample sizes, parameter settings and length of time aggregation. Both weighted and unweighted mean point estimates are close to the population values. The variability of the point estimates, as measured by the SDs, almost always falls, hence efficiency increases, on moving from unweighted GMM to weighted GMM, and then across the iterations. Mean SEs are in good agreement with the SDs (in case of eta some small discrepancies are due to a small number of extreme point estimates), so we expect inference to be reliable. Even sample sizes as low as 100 can be dealt with successfully, although variability of the estimates can be somewhat large; the variability falls substantially when sample sizes are increased from 100 to 500, which is still significantly smaller than in our empirical application. For samples of size 500 the inefficient unweighted estimator yields good and significant estimates across the two settings. Tables 1 also reveals that typically the precision of the estimates increases as the extent of time aggregation falls. We proceed to examine this more systematically.

We investigate the impact of the effect of time aggregation on the estimation's accuracy by varying the number of grid points – the smaller the number of grid point, the longer the intervals. More specifically, we consider individuals from time $t = 0.1$ to $t = 0.6$, and half this interval, and repeatedly half the resulting intervals. Let I_3 denote the case of two intervals, I_4 the case

	$[\eta, \sigma, s_0] = [1.5, .8, .5]$, I_1				$[\eta, \sigma, s_0] = [.2, .2236, .3536]$, I_2			
	iteration				iteration			
	0	1	3	$SE(\widehat{\theta}_3)$	0	1	3	$SE(\widehat{\theta}_3)$
$\widehat{\eta}$	1.468	1.504	1.489	0.146	0.285	0.197	0.193	0.058
(SD)	(0.526)	(0.267)	(0.232)	(0.036)	(0.088)	(0.100)	(0.108)	(0.018)
$\widehat{\sigma}$	0.785	0.792	0.789	0.026	0.247	0.221	0.223	0.005
(SD)	(0.190)	(0.039)	(0.034)	(0.003)	(0.088)	(0.008)	(0.006)	(0.000)
$\widehat{s}_0$	0.489	0.494	0.495	0.043	0.352	0.352	0.352	0.014
(SD)	(0.095)	(0.053)	(0.047)	(0.007)	(0.046)	(0.024)	(0.017)	(0.001)

Table 2: Simulation evidence II: Estimating the parameters of the intexpOU process.

of 4 intervals, and I_5 the case of 8 intervals. All intervals are regular. The mean estimates as well as their standard deviation (over all simulation runs) and their mean standard errors after iterating GMM three times are reported in table 3. The table indicates that the estimators are always unbiased, even in case I_1 where there are only two intervals. However, the interval length has a substantial impact on the standard errors. The larger the number of distinct intervals the smaller the standard error.

	$[\eta, \sigma, s_0] = [1.5, .8, .5]$			$[\eta, \sigma, s_0] = [.2, .2236, .3536]$		
	I_3	I_4	I_5	I_3	I_4	I_5
$\widehat{\eta}$	1.523	1.505	1.501	0.200	0.196	0.196
$(SD)\,,[SE]$	(.27),[.29]	(.18),[.18]	(.17),[.16]	(.06),[.06]	(.07),[.05]	(.07),[.04]
$\widehat{\sigma}$	0.800	0.798	0.799	0.222	0.222	0.223
$(SD)\,,[SE]$	(.07),[.08]	(.03),[.03]	(.02),[.02]	(.01),[.01]	(.01),[.01]	(.01),[.00]
$\widehat{s}_0$	0.500	0.497	0.498	0.353	0.352	0.351
$(SD)\,,[SE]$	(.04),[.04]	(.03),[.03]	(.03),[.03]	(.02),[.02]	(.02),[.01]	(.02),[.01]

Table 3: Simulation evidence: The effect of time agggregation.

In our final set of experiments we turn from the baseline intexpOU process to the more general Box Cox transformed model. The econometrician observes the intinvBCOU process $S_k^{(\lambda)}$. We reconsider the previous settings, and consider three cases for the Box Cox parameter λ, namely $\lambda = 0.5$, $\lambda = 0$, and $\lambda = -0.1$.

Table 4 reports the results. For $\lambda = 0.5$, all the estimates are very good across the investigated settings. The SDs diminish across the iterations, and the average SEs are in good agreement with the SDs. For $\lambda \leq 0$ the weighted estimation yielded a small number of large deviations, which

	$[\eta, \sigma, s_0] = [1.5, .8, .5],\ I_1$				$[\eta, \sigma, s_0] = [.2, .2236, .3536],\ I_2$			
	iteration				iteration			
	0	1	3	$SE(\widehat{\theta}_3)$	0	1	3	$SE(\widehat{\theta}_3)$
	$\lambda = 0.5$				$\lambda = 0.5$			
$\widehat{\eta}$	1.461	1.538	1.512	0.196	0.273	0.205	0.198	0.075
(SD)	(0.237)	(0.214)	(0.192)	(0.035)	(0.069)	(0.194)	(0.108)	(0.017)
$\widehat{\sigma}$	0.778	0.791	0.788	0.037	0.241	0.222	0.221	0.009
(SD)	(0.049)	(0.041)	(0.041)	(0.015)	(0.076)	(0.013)	(0.013)	(0.005)
$\widehat{s}_0$	0.488	0.492	0.491	0.060	0.351	0.351	0.350	0.029
(SD)	(0.094)	(0.060)	(0.057)	(0.005)	(0.018)	(0.030)	(0.016)	(0.008)
$\widehat{\lambda}$	0.465	0.465	0.460	0.219	0.500	0.551	0.515	0.302
(SD)	(0.192)	(0.215)	(0.205)	(0.040)	(0.269)	(0.294)	(0.265)	(0.053)
	$\lambda = 0$				$\lambda = 0$			
$\widehat{\eta}$	1.512	1.881	1.751	0.230	0.264	0.190	0.193	0.085
(SD)	(0.233)	(0.756)	(0.627)	(0.086)	(0.049)	(0.187)	(0.114)	(0.011)
$\widehat{\sigma}$	0.821	0.906	0.889	0.083	0.255	0.229	0.227	0.019
(SD)	(0.062)	(0.205)	(0.205)	(0.030)	(0.056)	(0.017)	(0.017)	(0.004)
$\widehat{s}_0$	0.521	0.575	0.560	0.074	0.357	0.360	0.359	0.043
(SD)	(0.087)	(0.164)	(0.132)	(0.024)	(0.021)	(0.033)	(0.022)	(0.006)
$\widehat{\lambda}$	0.040	0.043	0.066	0.256	0.027	0.105	0.052	0.382
(SD)	(0.187)	(0.251)	(0.258)	(0.047)	(0.283)	(0.325)	(0.307)	(0.060)
	$\lambda = -0.1$				$\lambda = -0.1$			
$\widehat{\eta}$	1.509	1.965	1.846	0.229	0.265	0.200	0.200	0.083
(SD)	(0.240)	(0.909)	(0.693)	(0.097)	(0.042)	(0.232)	(0.124)	(0.012)
$\widehat{\sigma}$	0.859	0.964	0.948	0.096	0.250	0.231	0.229	0.020
(SD)	(0.075)	(0.252)	(0.254)	(0.045)	(0.055)	(0.021)	(0.020)	(0.003)
$\widehat{s}_0$	0.539	0.595	0.577	0.079	0.363	0.366	0.364	0.045
(SD)	(0.092)	(0.219)	(0.170)	(0.037)	(0.022)	(0.042)	(0.028)	(0.005)
$\widehat{\lambda}$	0.004	0.012	0.001	0.254	-0.058	0.039	-0.043	0.393
(SD)	(0.172)	(0.287)	(0.356)	(0.076)	(0.268)	(0.460)	(0.338)	(0.065)

Table 4: Simulation evidence: Estimating the parameters of the Box Cox transformed model. Notes: As for Table 2. $N = 500$.

explains the occasional discrepancy between SDs and average SEs for iteration 3. Relatedly, the unweighted estimator is preferred to the weighted estimator in the first setting and for $\lambda = -0.1$ the mean estimate fails to pick up the sign of λ as the mean estimate is statistically insignificant, but recall that time aggregation in this setting is substantial. By contrast, the mean estimate of λ correctly picks up the sign in the second setting. The results for $\lambda = 0$ should also be compared to the results for the intexpOU process of Table 2. It is clear that the combination of an unconstrained estimation of λ and the use of fourth order approximations leads to no significant deterioration of the estimates. In summary, for sample sizes of 500, efficiency gains from iterated estimation only arise for $\lambda = 0.5$, the unweighted estimator yields typically good results, and the nested intexpOU population case is well estimated by the unconstrained estimator.

7 Empirical Application: Income Dynamics in the US.

We estimate the parameters of the structural continuous-time model using individual annual earnings data from the Panel Study of Income Dynamics (PSID).

7.1 The Data

The PSID data are provided in a standardized and easily accessible form by the Cross National Equivalent Files (CNEF) project (Frick et al., 2007).[9] We use a balanced panel of annual observations from 1989 until 1996, a total of 8 waves. $S_{k,i}$ are individual's i labor earnings in year k. These labor earnings include wages and salary from all employment including self-employment as well as bonuses, overtime and commissions. Annual earnings are measured in current 10,000 US dollars.

Turning to sample selection, we consider full-time employees having worked at least 1,500 hours per year. We follow standard practice and remove extreme outliers that could influence the estimation results by deleting the top and bottom 0.5% percent of observations in each wave. Our Mincerian covariates include sex, age, age^2, and years of education interpreted a (pre-determined) measure of human capital. In order to capture the effect of economic growth, we also include a time trend. Cohort effects are not identified when age and time are regressors present (Deaton and Paxson, 1994).

Table 5 reports some cross-sectional summary statistics of income. The number of persons in the panel is $N = 1,772$. The average age in the first wave is 36.4. Since the panel is balanced, average age increases by 1 each year.

[9] http://www.human.cornell.edu/che/PAM/Research/Centers-Programs/German-Panel/cnef.cfm.

Year	mean income [10^4\$]	st dev income [10^4\$]	Covariates	
1989	3.0430	1.7810	proportion male	74.5
1990	3.2560	1.8523	mean years of education	13.5
1991	3.4567	1.9448	mean age in first wave	36.4
1992	3.6423	2.0626		
1993	3.9523	2.2847		
1994	4.0787	2.4780		
1995	4.3003	2.6038	number of individuals	1,772
1996	4.4402	2.7473	number of observations	14,176

Table 5: PSID descriptive statistics.

Estimating a standard discrete-time random effects panel regression of log-earnings on covariates yields the following coefficient estimates (standard errors): sex: -0.319 (0.022), years of schooling: 0.114 (0.005), age: 0.081 (0.005), age squared: -0.00085 (5.3×10^{-5}) and linear time trend: 0.038 (0.002). The variance of the individual effect is estimated as 0.153 while the variance of the idiosyncratic error term is 0.078. The principal interest, however, is the stochastic structure of the continuous-time process $\{u(t)\}$.

7.2 Estimation Details

We estimate the model parameters by iterated GMM using the estimating equations described in Section 5. The length of vector f, i.e. the number of moment conditions used to identify the parameters, is $K(P+1) = 8 \cdot 6 = 48$.[10] We impose the stationarity restriction $a = 0$, i.e. the variance s_0^2 of the initial deviation $u(t_0)$ is not estimated but calculated from the estimates of η and σ^2 using (6). This is a mild restriction since the impact of the initial deviation vanishes over time if the OU process is stable.

7.3 Empirical Results

The estimates of the structural parameters are reported in Table 6. All parameters are statistically significant. The parameter estimates of β and γ are very plausible, close to the ones

[10] While there are six covariates (sex, education, age, age^2, time trend, and a constant for the intercept) present, we have to drop either the time trend or age from the list of covariates if we make the simplifying assumption in equation (15) that Z_2 does not change during the interval.

	parameter	estimate	SE	BC model	SE
intercept	μ	-2.3296	0.2315	-1.4072	0.2407
sex	β_1	-0.3038	0.0583	-0.2281	0.0535
education	β_2	0.1254	0.0216	0.1084	0.0153
age	γ_1	0.0928	0.0063	0.0654	0.0063
age^2	γ_2	-0.0010	0.0001	-0.0009	0.0001
time effect	γ_3	0.0407	0.0032	0.0360	0.0041
individual effect	σ_ε^2	0.1177	0.0125	0.0870	0.0225
OU parameter	η	2.3240	0.6475	1.955	0.4835
diffusion coefficient	σ^2	0.5033	0.1990	0.3287	0.1286
Box Cox parameter	λ	-	-	-0.1098	0.0576

Table 6: Coefficient estimates of the intexpOU earnings process estimated using PSID data.

estimated in a naive panel regression (though the standard errors are considerably larger since the naive panel model neglects the intertemporal dependence), and are of the same magnitudes as reported in the literature. This should not be surprising given the result stated in Lemma 6. The implied estimated variance of the initial deviation equals the unconditional variance $\hat{s}_0^2 = \hat{\sigma}^2/(2\hat{\eta}) = 0.1083$. As regards the interpretation of the parameters of the OU process, this is perhaps best done visually in terms of their implications on moments and the cross-sectional income distribution. Figure 1 (a) and (b) displays the goodness of fit of the model by comparing empirical sample moments of income and the moments predicted by the estimated model. It is evident that the moments are well estimated.

The estimates of the parameters of the Ornstein-Uhlenbeck component, $\hat{\eta}$ and $\hat{\sigma}^2$, imply that, given the individual effect ε_i, deviations from the expected log-income path decline relatively rapidly: The percentage deviation of the latent continuous-time income Y_{it} from its mean is expected to halve in about four months. In order to gauge the relative influence of the individual effect on the distribution we compute the cross-sectional coefficient of variation of the continuous-time process,

$$CV\left(Y_{it}\right) = \frac{\sqrt{Var\left(Y_{it}\right)}}{E\left(Y_{it}\right)} = \sqrt{\left(\exp\left(\sigma_\varepsilon^2 + \frac{\sigma^2}{2\eta}\right) - 1\right)}.$$

Inserting our parameter estimates yields $\widehat{CV} = 0.5035$. Setting $\sigma_\varepsilon^2 = 0$ to eliminate the individual effects the coefficient of variation decreases to 0.3382, while setting $\sigma^2 = 0$ to eliminate the error process results in $\widehat{CV} = 0.3534$. Even though the two components are not additive we can conclude that the influence of the individual effect and the error process on the distribution is of roughly the same order of magnitude.

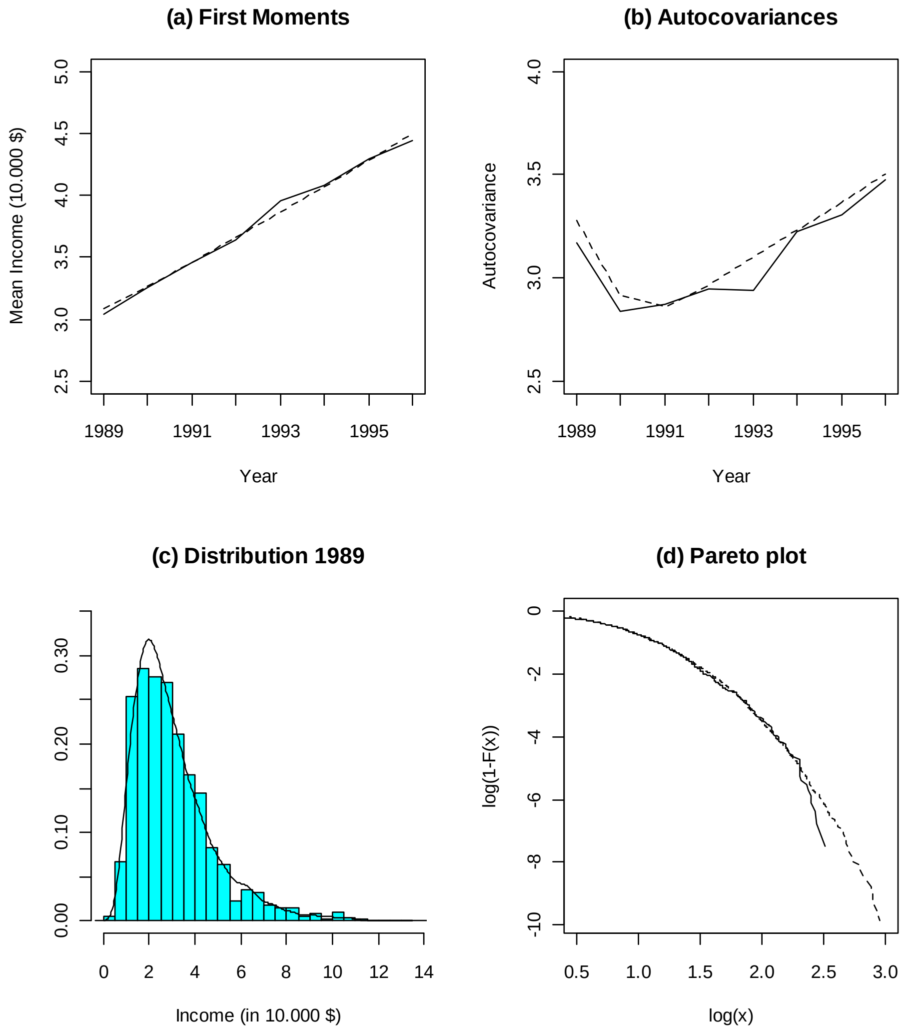

Figure 1: (a) Empirical means (solid line) and model implied first moments (dashed line); (b) empirical (solid) and theoretical (dashed) variances in 1989 and autocovariances between 1989 and 1990, ..., 1996; (c) histogram and model implied density of income 1989; (d) empirical (solid) and theoretical (dashed) Pareto plots of income 1989

We proceed to consider the entire income distribution. Lemma 2 states that the income density implied by the structural model is not tractable analytically. We therefore estimate the income density by Monte Carlo methods. In particular, we simulate income paths of $B = 20,000$ individuals having the same distribution of covariates as the original data. Figure 1 (c) depicts the histogram of the actual incomes for the year 1989, and the kernel density estimate of the simulated incomes. The model-based simulated density describes the actual income density well.

Finally, Figure 1 (d) shows the Pareto plots of the 1989 income data and the corresponding Pareto plot of the simulated incomes. Recall that a Pareto plot depicts $\log(1 - F(x))$ versus $\log(x)$. If the distribution is heavy-tailed, i.e. $1 - F(x) = x^{-1/\gamma} L_0(x)$ for sufficiently large x where L_0 is a slowly varying function and $\gamma > 0$, then the plot is a straight line with slope $-1/\gamma$ for sufficiently large x. As regards the theoretical Pareto plot, we know from Lemma 4 that the intexpOU process cannot generate heavy tails. This manifests itself in the curvature of the theoretical Pareto plot in the extreme right tail. However, Figure 1 (d) reveals that the empirical Pareto plot becomes a straight line in the rightmost tail of the income distribution. In particular, by inspection, the plot suggests the estimate $1/\widehat{\gamma} \approx -11$.

In the light of this tails behavior, we estimate the generalized model for the intinvBCOU process, in which the estimated Box Cox parameter, if negative, picks up the heavy tail. Table 6 reports the results. The Box Cox parameter is indeed estimated to be negative, but is small in magnitude and only marginally significant. Lemma 7 and the Remark following its proof in the Appendix state that $\gamma = |\lambda|^{-1}$ when $\lambda < 0$. The point estimate $\widehat{\lambda} = -.1$ suggests an estimate of γ which is very close to the slope of the Pareto plot in its right tail obtained by inspection. Turning to the the remaining covariate coefficients reported in the Table, these hardly change, while the estimates of the intercept μ and the OU drift parameter η have fallen in magnitude.

We summarize all this evidence by concluding that the estimated structural model of the intexpOU describes the empirical US earnings distributions and the intertemporal dependencies very well. Like all standard earnings models in the literature, the process generates a right tail of the earnings distribution which decays too fast, but this tail behavior is captured by the computationally more intensive generalized model of the intinvBCOU process. However, the difference in tail decay between the two models is not “too large”, so that for most applications the simpler intexpOU process provides a good description of all the features of the US earnings distribution.

8 Conclusion

We have considered continuous-time earnings models and their associated observable, time aggregated or 'integrated', processes. We have shown that time aggregation alters important statistical properties, for instance the integrated process does not inherit the lognormality and Markovianess of the underlying continuous-time process. The parameters of this process are estimable by GMM, and the finite sample performance of the estimator is shown to be good in several simulation studies. When applied to US panel data, the estimated models replicate well all important features of the actual earnings distribution. While the computationally more demanding intinvBCOU process does capture the heavy-tailedness of the actual earnings distribution, the estimate of the Box Cox parameter also reveals that the discrepancies between the speed of tail decays is relatively small. Hence we conclude that the simpler intexpOU process does a good job in describing the actual US earnings distributions.

References

Abowd, J. and D. Card (1989), "On the covariance structure of earnings and hours changes", *Econometrica*, 57, 411-445.

Baker, M. (1997), "Growth-rate heterogeneity and the covariance structure of life-cycle earnings", *Journal of Labor Economics*, 15, 338-375.

Barndorff-Nielsen, O. E. and N. Shepard (2001), "Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial econometrics (with discussion)", *J. Roy. Statist. Soc. Ser. B,* 63, 167-241.

Bhattacharya, R., E. Thomann, and E. Waymire (2001), "A note on the distribution of integrals of geometric Brownian motion", *Statistics and Probability Letters,* 55, 187-192.

Bick, B., H. Kraft and C. Munk (2009), "Investment, income, and incompleteness", working paper.

Bodie, Z., Merton, R.C. and W.F. Samuelson (1992), "Labor supply flexibility and portfolio choice in a life-cycle model", *Journal of Economic Dynamic and Control*, 16(3-4), 427-449.

Carr, P. and M. Schröder (2004), "Bessel processes, the integral of geometric Brownian motion, and Asian options", *Theory of Probability and Its Applications*, 48, 400-425.

Christiano, L. J., M. Eichenbaum, and D. Marshall (1991), "The permanent income hypothesis revisited", *Econometrica*, 59, 397-423.

Comte, F., V. Genon-Catalot, and Y. Rozenholc (2009), "Nonparametric adaptive estimation for integrated diffusions", *Stochastic Processes and their Applications*, 119, 811-834.

Deaton, A. and C. Paxson (1994), "Intertemporal choice and inequality", *Journal of Political Economy*, 102, 437-467.

Ditlevsen, S. and M. Sørensen (2004), "Inference for observations of integrated diffucion processes", *Scandinavian Journal of Statistics*, 31, 417-429.

Frick, J.R., S.P. Jenkins, D.R. Lillard, O. Lipps and M. Wooden (2007), "The Cross-National Equivalent File (CNEF) and its Member Country Household Panel Studies", *Schmollers Jahrbuch,* 127, 627-654.

Gloter, A. (2001), "Parameter estimation for a discrete sampling of an integrated Ornstein-Uhlenbeck process", *Statistics*, 35, 225-243.

Gloter, A. (2006), "Parameter estimation for a discretely observed integrated diffusion process", *Scandinavian Journal of Statistics*, 33, 83-104.

Guvenen, F. (2009), "An empirical investigation of labor income processes", *Review of Economic Dynamics*, 12, 58-79.

Harvey, A. C. and J. H. Stock (1989), "Estimating integrated higher-order continuous time autoregressions with an application to money-income causality", *Journal of Econometrics*, 42, 319-336.

Henderson, V. (2005), "Explicit solutions to an optimal portfolio choice problem with stochastic income", *Journal of Economic Dynamics and Control*, 29, 1237-1266.

Kaufmann, E. (2000), "Penultimate Approximations in Extreme Value Theory", *Extremes*, 3, 39-55.

Koo, H.K. (1998), "Consumption and portfolio selection with labor income: a continuous time approach", *Mathematical Finance*, 8, 49-65.

MaCurdy, T. (1982), "The use of time series processes to model the error structure of earnings in longitudinal data analysis", *Journal of Econometrics*, 18, 83-114.

Munk, C. and C. Sørensen (2009), "Dynamic asset allocation with stochastic income and interest rates", working paper.

Newey, W.K. and D. McFadden (1994), "Large sample estimation and hypothesis testing", in: *Handbook of Econometrics*, chapter 36.

Pareto, V. (1896), "La courbe de la répartition de la richesse", reprinted 1965 in G. Busoni (ed.): *Oeuvres completés de Vilfredo Pareto*, Tome 3: Écrits sur la courbe de la répartition de la richesse, Geneva: Librairie Droz. English translation in *Rivista di Politica Economica*, 87 (1997), 647-700.

Schluter, C. and M. Trede (2002), "Tails of Lorenz curves", *Journal of Econometrics*, 109, 151-166.

Schluter, C. and M. Trede (2008), "Identifying multiple outliers in heavy-tailed distributions with an application to market crashes", *Journal of Empirical Finance*, 15, 700-713.

Wang, N., (2004), "Precautionary saving with partially observed income", *Journal of Monetary Economics*, 51, 1645-1681.

Wang, N. (2006), "Generalizing the permanent-income hypothesis: revisiting Friedman's conjecture on consumption", *Journal of Monetary Economics*, 53, 737-752.

Wang, N. (2009), "Optimal consumption and asset allocation with unknown income growth", *Journal of Monetary Economics*, 56, 524-534.

A Moment Expressions for the intexpOU Process with $\tilde{y}(t) \equiv 0$

The presence of covariates in the $\tilde{y}(t)$ process precludes us from saying much analytically, but in the absence of the $\tilde{y}(t)$ process we can obtain some useful insights about the moments of S_k. For expositional brevity we focus on regular intervals of length Δ. For the first moment we obtain an exact result. Exact results are also available in the stationary case (when $a = 0$), otherwise we state approximations for variances and covariances.

Lemma 8 *Consider the case when $\tilde{y}(t) \equiv 0$ and regular intervals of length Δ. Then*

$$E(S_k) \times \exp\left(-\frac{\sigma^2}{4\eta}\right) = \Delta + \frac{1}{2\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{a}{2}e^{-2\eta(t_{k-1})\Delta}\right]^i\left[1 - e^{-2i\eta\Delta}\right].$$

In the stationary case with $a = 0$, or asymptotically with $k \to \infty$, we have $E(S_k) = \Delta \exp\left(\sigma^2/4\eta\right)$, while the expectation becomes unbounded for $\Delta \to \infty$.

Exact expressions for covariances and variances are only available in the stationary case considered explicitly below in Lemma 10. Asymptotics for $k \to \infty$ i.e. $t_{k-1} \to \infty$ yield simple results and are stated next, while approximations for the general case are stated below in Lemma 12.

A.1 Moments' Asymptotics

Lemma 9 *Consider the case when $\tilde{y}(t) \equiv 0$, let $\eta > 0$ and consider regular intervals of fixed length Δ. As $k \to \infty$*

$$\begin{aligned}
E\{S_k\} &\to \exp\left(\frac{\sigma^2}{4\eta}\right)\Delta, \\
Cov\{S_l, S_k\}_{l\ fixed} &\to 0, \\
Var\{S_k\} \times \exp\left(-\frac{\sigma^2}{2\eta}\right)\frac{1}{2} &\to \frac{1}{\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{\sigma^2}{2\eta}\right]^i\left(\Delta + \frac{1}{i\eta}\left[e^{-i\eta\Delta} - 1\right]\right).
\end{aligned}$$

A.2 The Exact Moments of $\{S_k\}$ in the Stationary Case

In the stationary case $a = 0$, and we can state exact expressions for the covariance structure:

Lemma 10 *Consider the case when $\tilde{y}(t) \equiv 0$, let $a = 0$ and consider regular intervals of fixed* *length Δ. Then*

$$\begin{aligned}
E\{S_k\} &= \exp\left(\frac{\sigma^2}{4\eta}\right)\Delta \\
Cov\{S_l,S_k\}_{l<k}\times\exp\left(-\frac{\sigma^2}{2\eta}\right) &= \frac{1}{\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{\sigma^2}{2\eta}\right]^i\frac{1}{i\eta}e^{-i\eta\Delta(k-l)}\times \\
&\quad \left[e^{i\eta\Delta}-1\right]\left[1-e^{-i\eta\Delta}\right] \\
Var\{S_k\}\times\exp\left(-\frac{\sigma^2}{2\eta}\right)\frac{1}{2} &= \frac{1}{\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{\sigma^2}{2\eta}\right]^i\left(\Delta+\frac{1}{i\eta}\left[e^{-i\eta\Delta}-1\right]\right).
\end{aligned}$$

Of course Lemma 10 specializes to Lemma 9 as $k\to\infty$.

Moreover, we can derive the spectral density for the stationary process.

Lemma 11 *Consider the case when $\tilde{y}(t)\equiv 0$, let $a=0$ and consider regular intervals of fixed length Δ. The spectral density is*

$$f_S(\lambda)=\sum_{j=1}^{\infty}\frac{c_j}{2\pi}\left[\frac{1-e^{-2j\eta\Delta}}{1-2e^{-j\eta\Delta}\cos\lambda+e^{-2j\eta\Delta}}\right] \tag{16}$$

with $c_j=\exp\left(\frac{\sigma^2}{2\eta}\right)\frac{1}{j}\frac{1}{j!}\left[\frac{\sigma^2}{2\eta}\right]^j\frac{1}{j\eta^2}\left[e^{j\eta\Delta}+e^{-j\eta\Delta}-2\right]$ and $\lambda\in[-\pi,\pi]$.

The first order term in the series of $f_S(\lambda)$, $\frac{c_1}{2\pi}\left[1-2e^{-\eta\Delta}\cos\lambda+e^{-2\eta\Delta}\right]^{-1}$, is the spectral density of an AR(1) process with coefficient $e^{-\eta\Delta}$, with the variance of the white noise error equal to c_1 and depending on the autoregressive coefficient.[11]

A.3 Covariance Approximations for the General Case

The exact expression for the first moment is stated above in Lemma 8. Exact expressions for the covariance structure are not available in the general case, but we can state the following approximations:

Lemma 12 *Consider the case when $\tilde{y}(t)\equiv 0$ and assume that $\eta>0$. Then*

$$\begin{aligned}
&Cov(S_l,S_k)_{l<k}\times\exp\left(-\frac{\sigma^2}{2\eta}\right) \\
\simeq\ & \frac{1}{\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{\sigma^2}{2\eta}\right]^i\frac{1}{i\eta}\left[e^{i\eta\Delta}-1\right]\left[1-e^{-i\eta\Delta}\right]e^{-i\eta(k-l)\Delta}- \\
& \frac{1}{2\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{a}{2}\right]^i\frac{1}{2i\eta}e^{-2i\eta(l-1)\Delta}\left[e^{-2i\eta\Delta}-1\right]^2e^{-2i\eta(k-1)\Delta}.
\end{aligned}$$

[11] For completeness we also state the spectral density of the underlying continuous process as $f_Y(\lambda)=\frac{1}{2\pi}\sigma^2\exp\left\{\frac{\sigma^2}{2\eta}\right\}\sum_{j=1}^{\infty}\frac{1}{j!}\left(\frac{\sigma^2}{2\eta}\right)^j 2j\eta\frac{1}{j^2\eta^2+\lambda^2}$ for $\lambda\in R$.

$$Var\left\{S_k\right\}\times\exp\left(-\frac{\sigma^2}{2\eta}\right) \simeq \frac{2}{\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{\sigma^2}{2\eta}\right]^i\left(\Delta+\frac{1}{i\eta}\left[e^{-i\eta\Delta}-1\right]\right)- \frac{1}{2\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{a}{2}\right]^i\frac{1}{2i\eta}\left[e^{-2i\eta\Delta}-1\right]^2e^{-4i\eta(k-1)\Delta}.$$

The order of the approximation is $a\Delta\frac{1}{\eta}e^{-\eta(k-1)\Delta}\left[1-e^{-\eta\Delta}\right]$.

In the stationary case $a=0$ and this covariance and variance immediately simplify to the statement in Lemma 10, while with t_{l-1} fixed and $t_{k-1}\to\infty$ we immediately obtain the statement of Lemma 9. Similar approximations can be derived for the case $\eta<0$.

B Proofs

We briefly state two Lemmas which will be used in the subsequent proofs.

Lemma 13 *Let* $X=(X_1,\ldots,X_T)^\top$ *be a multivariate normal random variable. Its moment generating function is*

$$M_X(\theta)\equiv E\left(e^{\theta^\top X}\right)=\exp\left(\mu^\top\theta+\frac{1}{2}\theta^\top\Sigma\theta\right)$$

where μ *is the expectation vector and* $\Sigma=[\sigma_{st}]$ *is the covariance matrix. The expectation of the exponentiated sum is then*

$$E\left(e^{X_1+\ldots+X_K}\right)=M_X(1)=\exp\left(\sum_{k=1}^{K}\mu_k+\frac{1}{2}\sum_{k=1}^{K}\sum_{r=1}^{K}\sigma_{kr}\right), \tag{17}$$

and $Y=\exp(X)$ *is multivariate lognormal with expectations and covariances given by*

$$\begin{aligned} E(Y_k) &= \exp\left(\mu_k+\frac{1}{2}\sigma_{kk}\right) \\ Cov(Y_k,Y_r) &= \exp\left((\mu_k+\mu_r)+\frac{1}{2}(\sigma_{kk}+\sigma_{rr})\right)(\exp(\sigma_{kr})-1). \end{aligned}$$

Direct computations lead to the following lemma.

Lemma 14 *The process* $\{\tilde{y}(t):t\geq t_0\}$ *is Gaussian with*

$$\begin{aligned} E(\tilde{y}(t)) &= \mu+Z_1^\top\beta+Z_{2,t}^\top\gamma \\ Cov(\tilde{y}(s),\tilde{y}(t)) &= \sigma_\varepsilon^2, \end{aligned}$$

and the process $\{u(t):t\geq t_0\}$ *is Gaussian with*

$$\begin{aligned} E(u(t)) &= 0 \\ Cov(u(s),u(t)) &= \left[s_0^2+\frac{\sigma^2}{2\eta}\left(e^{2\eta\min(s,t)}-1\right)\right]e^{-\eta(t+s)} \\ &= \left[\frac{\sigma^2}{2\eta}e^{-\eta(t-s)}+e^{-\eta(t+s)}\left(s_0^2-\frac{\sigma^2}{2\eta}\right)\right]_{s\leq t}. \end{aligned}$$

Proof of Proposition 1. By direct computation using the Lemmas 13 and 14. In particular $\{Y(t) : t \geq t_0\}$ is a log-normal process. Finally, $\{\exp(u(t))\}$ is Markov since $u(t)$ is, because the increments are independent, and $\{Y(t)\}$ is Markov given the independence of $\{\tilde{y}(t)\}$ and $\{u(t)\}$. ■

Proof of Lemma 2. The distribution of S_k has no closed form expression, since the distribution of the sum of dependent log-normal variates has no exact closed form. As regards the non-Markov property, only consider the OU process and intervals of equal length $0 = t_0 < t_1 = \Delta < t_2 = 2\Delta < \ldots < t_T = T\Delta$. Using (4) we find

$$\begin{aligned} S_k &= \int_{(k-1)\Delta}^{k\Delta} \exp(u(t))\, dt \\ &= \int_0^{\Delta} \exp\left(u((k-1)\Delta)\, e^{-\mu t} + \sigma \int_{(k-1)\Delta}^{(k-1)\Delta+t} e^{\mu(s-t)} dW_s \right) dt. \end{aligned}$$

If the value of the latent process at the beginning of the interval, $u((k-1)\Delta)$, was known, S_k would not depend on $S_{k-1}, S_{k-2}, \ldots$ However, since the process $u(t)$ is latent we can only use the conditional distribution of $u((k-1)\Delta)$ given $S_{k-1}, S_{k-2}, \ldots$ to determine the distribution of S_k. As the same holds true for S_{k-1} and $u((k-2)\Delta)$, the process $\{S_k : k = 1, \ldots, T\}$ cannot be Markov. ■

Proof of Lemma 7. We recall some results from extreme value theory (EVT). The Generalized Extreme Value distribution is given by

$$G(x) = \exp\left(-\left[1 + \frac{\gamma}{\sigma}(x-\mu)_+^{-1/\gamma} \right] \right)$$

where γ is the shape parameter of interest. It is the limit distribution of the normalized maximum of a random variable X with distribution F_X (if the limit exists). $\gamma > 0$ is the Frechet case, so F_X is heavy tailed, $\gamma \to 0$ is the Gumbel case so F_X is slowly varying in its right tail, and $\gamma < 0$ is the negative Weibull case. Let x_F be the upper end point of the distribution F_X, and let $h_X(x) = \frac{1-F_X(x)}{f_X(x)}$ be the reciprocal of the hazard function. A well-known result from EVT (Kaufmann, 2000) states that $\gamma_X = \lim_{x \to x_F} h'_X(x)$.

Consider the Box Cox transformation $g(x) = \frac{x^\lambda - 1}{\lambda}$ with $x > 0$ and $\lambda \neq 0$. For later reference we have $g'(x) = x^{\lambda-1}$ and $g''(x) = (\lambda-1)x^{\lambda-2}$ so $g''(x) / [g'(x)]^2 = (\lambda-1)x^{-\lambda}$. Let $Y = g^{-1}(U)$. We prove the result stated in the Lemma more generally for distributions of U which lie in specific domains of attractions which depend on the sign of λ.

We have

$$F_Y(y) = \Pr\{Y \leq y\} = \Pr\left\{g^{-1}(U) \leq y\right\} = \Pr\{U \leq g(y)\} = F_U(g(y)). \tag{18}$$

The associated density is $f_Y(y) = f_U(g(y))\, g'(y)$, and

$$\begin{aligned} h_Y(y) &= \frac{1 - F_Y(y)}{f_Y(y)} = \frac{1 - F_U(g(y))}{f_U(g(y))} \frac{1}{g'(y)} \\ &= h_U(g(y)) \frac{1}{g'(y)}. \end{aligned}$$

Differentiating the latter yields

$$h'_Y(y) = h'_U(g(y)) - h_U(g(y)) \frac{g''(y)}{[g'(y)]^2}.$$

In the particular case of the Box Cox transform, $g''(y) / [g'(y)]^2 = (\lambda - 1)\, y^{-\lambda}$, so

$$\gamma_Y = \gamma_U - (\lambda - 1) \lim_{y \to y_F} h_U(g(y))\, y^{-\lambda}. \tag{19}$$

As regards the Box Cox parameter λ, we distinguish between three cases.

Case (a) If $\lambda > 0$, assume that the distribution of U is in the domain of attraction of the Gumbel distribution, so that

$$h_U(x) = Cx^{-\alpha} \left(1 + O\left(x^{-\varepsilon}\right)\right)$$

with $\alpha > -1$, $\varepsilon > 0$ and $C > 0$. This class includes the normal distribution with $\alpha = 1$ and the lognormal distribution with $\alpha = 0$. We have $y_F = \infty$. Then $h'_u(x) = -\alpha C x^{-(\alpha+1)} (1 + O(x^{-\varepsilon}))$, and thus $\gamma_U = \lim_{x \to \infty} h'_u(x) = 0$. Moreover

$$h_U(g(y))\, y^{-\lambda} = C\lambda^{-\alpha} \left[y^{\lambda} - 1\right]^{-\alpha} y^{-\lambda} \left(1 + O\left(\left[y^{\lambda} - 1\right]^{-\varepsilon}\right)\right).$$

Then $\lim_{y \to \infty} h_U(g(y))\, y^{-\lambda} = 0$ for $\alpha \geq 0$. We have the result that $\gamma_Y = 0$ so F_Y is slowly varying in its right tail.

Case (b) When $\lambda = 0$, $h'_Y(y) = h'_U(g(y)) + \lim_{y \to \infty} h_U(g(y))$, with $g(y) = \ln(y)$ so for $\alpha \geq 0$, $\gamma_Y = 0$ again.

Case (c) Next, consider $\lambda < 0$. Then g has an asymptote at $|\lambda|^{-1}$, so that U cannot exceed $|\lambda|^{-1}$. Assume that the distribution of U satisfies this restriction. In particular, the distribution of U could be normal but truncated at $u_F \to |\lambda|^{-1}$so that $y_F = g^{-1}(u_F)$. Therefore the distribution of U is now in the domain of attraction of the (negative) Weibull distribution, so that

$$1 - F_U(u) = C(u_F - u)^{\alpha} \left[1 + D(u_F - u)^{\beta} + O\left((u_F - u)^{\beta+\varepsilon}\right)\right] \tag{20}$$

with $\alpha, \beta, \varepsilon, C > 0$ and $D \in R$, and $u_F < \infty$. Then

$$f_U(u) = \alpha C(u_F - u)^{\alpha-1} \left[1 + ((\alpha + \beta)/\alpha)\, D(u_F - u)^{\beta} + O\left((u_F - u)^{\beta+\varepsilon}\right)\right]$$

and

$$h_U(u) = \alpha^{-1}(u_F - u)\frac{1 + D(u_F - u)^{\beta} + O\left((u_F - u)^{\beta+\varepsilon}\right)}{1 + ((\alpha+\beta)/\alpha)\,D(u_F - u)^{\beta} + O\left((u_F - u)^{\beta+\varepsilon}\right)}.$$

It follows that $\gamma_U = \lim_{u\to u_F} h_U^{'}(u) = -\alpha^{-1}$. The Box Cox transform is $g(x) = \frac{1-x^{-|\lambda|}}{|\lambda|}$ and $g''(y)/[g'(y)]^2 = -(1+|\lambda|)\,y^{|\lambda|}$, so

$$\gamma_Y = \gamma_U + (1+|\lambda|)\lim_{y\to y_F} h_U(g(y))\,y^{|\lambda|}.$$

We have for $u_F \to |\lambda|^{-1}$ and thus $y_F \to \infty$

$$\begin{aligned}
\lim_{y\to\infty} h_U(g(y))\,y^{|\lambda|} &= \lim_{y\to\infty} \alpha^{-1}\left(|\lambda|^{-1} - \frac{1-y^{-|\lambda|}}{|\lambda|}\right) y^{|\lambda|} \\
&= \lim_{y\to\infty} \alpha^{-1}\left(\frac{y^{-|\lambda|}}{|\lambda|}\right) y^{|\lambda|} \\
&= \alpha^{-1}|\lambda|^{-1}.
\end{aligned}$$

In summary

$$\begin{aligned}
\gamma_Y &= -\alpha^{-1} + (1+|\lambda|)\,\alpha^{-1}|\lambda|^{-1} \\
&= \alpha^{-1}|\lambda|^{-1}.
\end{aligned}$$

Therefore $\gamma_Y > 0$ and the distribution is heavy-tailed. ■

Remark 15 *If the distribution of U is a truncated normal distribution, then F_U is in the domain of attraction of the negative Weibull distribution given by (20) with $\alpha = 1$, and then $\gamma_Y = |\lambda|^{-1}$ if $\lambda < 0$.*

Proof: Consider

$$1 - F_U(u) = c\frac{1}{\sqrt{2\pi}\sigma}\int_0^{u_F-u} \exp\left(-\frac{1}{2}\left(\frac{(u-\mu)+\varepsilon}{\sigma}\right)^2\right) d\varepsilon.$$

The claim follows using a Taylor expansion about $\varepsilon = 0$. ■

For the next set of proofs, the following result is useful.

Lemma 16 *Define the error integral by*

$$EI(y;\widetilde{a}) \equiv \int \frac{\exp(\widetilde{a}y)}{y}dy = \log y + \frac{\widetilde{a}y}{1} + \frac{1}{2}\frac{(\widetilde{a}y)^2}{2!} + \frac{1}{3}\frac{(\widetilde{a}y)^3}{3!} + .. \tag{21}$$

Then

$$\begin{aligned}\int_{t_{k-1}}^{t_k} \exp\left(\widetilde{a}e^{bt}\right) dt &= \left(\frac{1}{b}\right)\int_{\tau_{k-1}}^{\tau_k} \frac{\exp\left(\widetilde{a}y\right)}{y} dy \\ &= \left(\frac{1}{b}\right) EI\left(y;\widetilde{a}\right)|_{e^{bt_{k-1}}}^{e^{bt_k}} \\ &= t_k - t_{k-1} + \frac{1}{b}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\widetilde{a}^i\left[e^{ibt_k} - e^{ibt_{k-1}}\right].\end{aligned}$$

The Lemma follows after the change of variables $y = e^{bt}$ with inverse Jacobian $\partial y/\partial t = by$.

Proof of Lemma 8. We have

$$\begin{aligned}E\left\{S_k\right\} &= \int_{\tau=0}^{\Delta} E\left\{Y\left(t_{k-1}+\tau\right)\right\} d\tau \\ &= \exp\left(\frac{\sigma^2}{4\eta}\right)\int_{\tau=0}^{\Delta}\exp\left(\frac{a}{2}e^{-2\eta t_{k-1}}e^{-2\eta\tau}\right) d\tau.\end{aligned}$$

Hence using Lemma 16 we obtain

$$E\left\{S_k\right\} \times \exp\left(-\frac{\sigma^2}{4\eta}\right) = \Delta - \frac{1}{2\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{a}{2}e^{-2\eta t_{k-1}}\right]^i\left[e^{-2\eta i\Delta} - 1\right].$$

■

Proof of Lemma 9.

(i) The expression for the first moment follows immediately from Lemma 8.

(ii) The covariances. Fix s and apply Proposition 1. Then

$$\lim_{t\to\infty} Cov\left(Y\left(s\right), Y\left(t\right)\right)_{s<t} \times \exp\left(-\frac{\sigma^2}{2\eta}\right) = 0$$

and therefore

$$\begin{aligned}&\lim_{t_{k-1}\to\infty} Cov\left\{S_l, S_k\right\}_{t_{l-1}\text{ fixed}} \\ =\; &\lim_{t\to\infty}\int_{s=t_{l-1}}^{t_{l-1}+\Delta}\left\{\int_{\tau=0}^{\Delta} Cov\left(Y\left(s\right), Y\left(t_{k-1}+\tau\right)\right) d\tau\right\} ds \\ =\; &0.\end{aligned}$$

(iii) The variances. Letting both $s, t \to \infty$ we have by Proposition 1

$$\begin{aligned}Cov\left(Y\left(s\right), Y\left(t\right)\right)_{s<t} \times \exp\left(-\frac{\sigma^2}{2\eta}\right) &\to \exp\left(\frac{\sigma^2}{2\eta}e^{-\eta(t-s)}\right) - 1, \\ Cov\left(Y\left(s\right), Y\left(t\right)\right)_{s>t} \times \exp\left(-\frac{\sigma^2}{2\eta}\right) &\to \exp\left(\frac{\sigma^2}{2\eta}e^{-\eta(s-t)}\right) - 1.\end{aligned}$$

Then

$$
\begin{aligned}
Var\left\{S_k\right\} &= \int_{t=t_{k-1}}^{t_k} \int_{s=t_{k-1}}^{t} Cov\left(Y\left(s\right), Y\left(t\right)\right) dsdt \\
&\quad + \int_{t=t_{k-1}}^{t_k} \int_{s=t}^{t_k} Cov\left(Y\left(s\right), Y\left(t\right)\right) dsdt
\end{aligned}
$$

By Lemma 16 with $\widetilde{a} = \left(\sigma^2/2\eta\right) e^{-\eta t}$ we have

$$
\begin{aligned}
\int_{s=t_{k-1}}^{t} \exp\left(\widetilde{a} e^{\eta s}\right) ds &= \left(t - t_{k-1}\right) + \frac{1}{\eta} \sum_{i=1}^{\infty} \frac{1}{i} \frac{1}{i!} \widetilde{a}^i \left(e^{i\eta t} - e^{i\eta t_{k-1}}\right) \\
&= \left(t - t_{k-1}\right) + \frac{1}{\eta} \sum_{i=1}^{\infty} \frac{1}{i} \frac{1}{i!} \left[\frac{\sigma^2}{2\eta}\right]^i \left(1 - e^{-i\eta t} e^{i\eta t_{k-1}}\right)
\end{aligned}
$$

Therefore integrating over t yields

$$
\begin{aligned}
& \exp\left(-\frac{\sigma^2}{2\eta}\right) \times \int_{t=t_{k-1}}^{t_k} \int_{s=t_{k-1}}^{t} Cov\left(Y\left(s\right), Y\left(t\right)\right) dsdt \\
= \;& \frac{\Delta}{\eta} \sum_{i=1}^{\infty} \frac{1}{i} \frac{1}{i!} \left[\frac{\sigma^2}{2\eta}\right]^i + \frac{1}{\eta^2} \sum_{i=1}^{\infty} \frac{1}{i^2} \frac{1}{i!} \left[\frac{\sigma^2}{2\eta}\right]^i \left[e^{-i\eta\Delta} - 1\right].
\end{aligned}
$$

A similar calculation for $\exp\left(-\sigma^2/2\eta\right) \times \int_{t=t_{k-1}}^{t_k} \int_{s=t}^{t_{k-1}} Cov\left(Y\left(s\right), Y\left(t\right)\right) dsdt$ yields the same result, so putting everything together, we have

$$
Var\left\{S_k\right\} \exp\left(-\frac{\sigma^2}{2\eta}\right) \frac{1}{2} \rightarrow \frac{\Delta}{\eta} \sum_{i=1}^{\infty} \frac{1}{i} \frac{1}{i!} \left[\frac{\sigma^2}{2\eta}\right]^i + \frac{1}{\eta^2} \sum_{i=1}^{\infty} \frac{1}{i^2} \frac{1}{i!} \left[\frac{\sigma^2}{2\eta}\right]^i \left[e^{-i\eta\Delta} - 1\right].
$$

■

Proof of Lemma 10. Use similar calculations as in the proof of Lemma 9. ■

Proof of Lemma 11. Using Lemma 10 with $k = l + h$ the autocovariance function is, for all h,

$$
\gamma_S\left(h\right) = \sum_{j=1}^{\infty} c_j e^{-bj|h|}
$$

with $b = \eta\Delta > 0$ and $c_j = \exp\left(\frac{\sigma^2}{2\eta}\right)\frac{1}{j}\frac{1}{j!}\left[\frac{\sigma^2}{2\eta}\right]^j \frac{1}{j\eta^2}\left[e^{j\eta\Delta} + e^{-j\eta\Delta} - 2\right]$. The spectral density is defined as $f_S(\lambda) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-i\lambda h}\gamma_S(h)$ for $i = \sqrt{-1}$. Consider first the first order term

$$\begin{aligned}
\widetilde{f}_S(\lambda) &= \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-i\lambda h} c_1 e^{-b|h|} \\
&= \frac{c_1}{2\pi}\left[1 + \sum_{h\geq 1}^{\infty}\left(e^{-(i\lambda+b)}\right)^h + \sum_{h\leq -1}^{-\infty}\left(e^{-(i\lambda-b)}\right)^h\right] \\
&= \frac{c_1}{2\pi}\left[1 + \sum_{h\geq 1}^{\infty}\left(e^{-(i\lambda+b)}\right)^h + \sum_{h\geq 1}^{\infty}\left(e^{(i\lambda-b)}\right)^h\right] \\
&= \frac{c_1}{2\pi}\left[1 + \frac{e^{-(i\lambda+b)}}{1 - e^{-(i\lambda+b)}} + \frac{e^{(i\lambda-b)}}{1 - e^{(i\lambda-b)}}\right] \\
&= \frac{c_1}{2\pi}\left[\frac{1 - e^{-2b}}{1 - e^{-b}\left(e^{i\lambda} + e^{-i\lambda}\right) + e^{-2b}}\right] \\
&= \frac{c_1}{2\pi}\left[\frac{1 - e^{-2b}}{1 - 2e^{-b}\cos\lambda + e^{-2b}}\right].
\end{aligned}$$

The spectral density is therefore

$$f_S(\lambda) = \sum_{j=1}^{\infty}\frac{c_j}{2\pi}\left[\frac{1 - e^{-2jb}}{1 - 2e^{-jb}\cos\lambda + e^{-2jb}}\right].$$

■

Remark: The spectral density of the underlying continuous-time process $Y(t)$ is

$$f_Y(\lambda) = \frac{1}{2\pi}\sigma^2 \exp\left\{\frac{\sigma^2}{2\eta}\right\}\sum_{j=1}^{\infty}\left(\frac{\sigma^2}{2\eta}\right)^j 2j\eta\frac{1}{j^2\eta^2 + \lambda^2}.$$

Proof. The autocovariance function is

$$\gamma_Y(h) = \exp\left\{\frac{\sigma^2}{2\eta}\right\}\left[\exp\left\{\frac{\sigma^2}{2\eta}e^{-\eta|h|}\right\} - 1\right].$$

Then

$$
\begin{aligned}
f_Y(\lambda) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\lambda t}\gamma_Y(t)\,dt \\
&= \frac{1}{2\pi}\exp\left\{\frac{\sigma^2}{2\eta}\right\}\int_{-\infty}^{\infty} e^{-i\lambda t}\left[\exp\left\{\frac{\sigma^2}{2\eta}e^{-\eta|t|}\right\}-1\right]dt \\
&= \frac{1}{2\pi}\exp\left\{\frac{\sigma^2}{2\eta}\right\}\int_{-\infty}^{\infty} e^{-i\lambda t}\sum_{j=1}^{\infty}\frac{1}{j!}\left(\frac{\sigma^2}{2\eta}\right)^j e^{-j\eta|t|}dt \\
&= \frac{1}{2\pi}\exp\left\{\frac{\sigma^2}{2\eta}\right\}\sum_{j=1}^{\infty}\frac{1}{j!}\left(\frac{\sigma^2}{2\eta}\right)^j \int_{-\infty}^{\infty} e^{-i\lambda t}e^{-j\eta|t|}dt \\
&= \frac{1}{2\pi}\exp\left\{\frac{\sigma^2}{2\eta}\right\}\sum_{j=1}^{\infty}\frac{1}{j!}\left(\frac{\sigma^2}{2\eta}\right)^j \left[\int_{0}^{\infty} e^{i\lambda t}e^{-j\eta t}dt+\int_{0}^{\infty} e^{-i\lambda t}e^{-j\eta t}dt\right] \\
&= \frac{1}{2\pi}\exp\left\{\frac{\sigma^2}{2\eta}\right\}\sum_{j=1}^{\infty}\frac{1}{j!}\left(\frac{\sigma^2}{2\eta}\right)^j \left[\frac{1}{j\eta+i\lambda}+\frac{1}{j\eta-i\lambda}\right] \\
&= \frac{1}{2\pi}\sigma^2\exp\left\{\frac{\sigma^2}{2\eta}\right\}\sum_{j=1}^{\infty}\frac{1}{j!}\left(\frac{\sigma^2}{2\eta}\right)^j 2j\eta\frac{1}{j^2\eta^2+\lambda^2}.
\end{aligned}
$$

■

Proof of Lemma 12. Consider the covariance expressions. We have

$$
\begin{aligned}
& Cov\left(S_l,S_k\right)_{l<k}\times\exp\left(-\frac{\sigma^2}{2\eta}\right) \\
= & \int_{t_{l-1}}^{t_l}\int_{t_{k-1}}^{t_k}\left[\exp\left(\frac{\sigma^2}{2\eta}e^{-\eta t}e^{\eta s}g(s,a)\right)-\exp\left(\frac{a}{2}e^{-2\eta t}e^{-2\eta s}\right)\right]dsdt
\end{aligned}
$$

with $g(s,a)=1+a\left(\frac{\sigma^2}{2\eta}\right)^{-1}e^{-2\eta s}+\frac{a}{2}\left(\frac{\sigma^2}{2\eta}\right)^{-1}e^{-\eta t}e^{-3\eta s}$. Hence $g(s,a)\simeq 1$ for $\eta>0$ and sufficiently large s or exactly for $a=0$. Using then Lemma 16 as usual we obtain

$$
\begin{aligned}
& Cov\left(S_l,S_k\right)_{l<k}\times\exp\left(-\frac{\sigma^2}{2\eta}\right) \\
\simeq & \frac{1}{\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{\sigma^2}{2\eta}\right]^i\frac{1}{i\eta}e^{i\eta t_{l-1}}\left[e^{i\eta\Delta}-1\right]\left[1-e^{-i\eta\Delta}\right]e^{-i\eta t_{k-1}}- \\
& \frac{1}{2\eta}\sum_{i=1}^{\infty}\frac{1}{i}\frac{1}{i!}\left[\frac{a}{2}\right]^i\frac{1}{2i\eta}e^{-2i\eta(t_{l-1}+t_{k-1})}\left[e^{-2i\eta\Delta}-1\right]^2.
\end{aligned}
$$

Next, consider the approximation error, which is the integral of

$$
\begin{aligned}
& \exp\left(\frac{\sigma^2}{2\eta}e^{-\eta t}e^{\eta s}\right)\left[\exp\left(ae^{-\eta t}e^{-\eta s}+\frac{a}{2}e^{-2\eta t}e^{-2\eta s}\right)-1\right] \\
= & \exp\left(\frac{\sigma^2}{2\eta}e^{-\eta t}e^{\eta s}\right)\times\left[ae^{-\eta t}e^{-\eta s}+o\left(ae^{-\eta t}e^{-\eta s}\right)\right].
\end{aligned}
$$

Hence integrating the leading term yields

$$
\begin{aligned}
& a \int_{t=t_{k-1}}^{t_k} e^{-\eta t} \int_{s=t_{l-1}}^{t_l} \exp\left(\frac{\sigma^2}{2\eta} e^{-\eta t} e^{\eta s} - \eta s \right) ds dt \\
< \; & a \int_{t=t_{k-1}}^{t_k} e^{-\eta t} \int_{s=t_{l-1}}^{t_l} \exp\left(\frac{\sigma^2}{2\eta} e^{-\eta t} e^{\eta s} \right) ds dt \\
= \; & a \int_{t=t_{k-1}}^{t_k} e^{-\eta t} \left[\Delta + \frac{1}{\eta} \sum_{i=1}^{\infty} \frac{1}{i} \frac{1}{i!} \left(\frac{\sigma^2}{2\eta} \right)^i e^{-i\eta t} \left[e^{i\eta t_l} - e^{i\eta t_{l-1}} \right] \right] \\
= \; & a\Delta \frac{1}{\eta} e^{-\eta(k-1)\Delta} \left[1 - e^{-\eta\Delta} \right] \\
& + a \frac{1}{\eta} \sum_{i=1}^{\infty} \frac{1}{i} \frac{1}{i!} \left(\frac{\sigma^2}{2\eta} \right)^i e^{i\eta(l-1)\Delta} \left[e^{i\eta\Delta} - 1 \right] \frac{1 - e^{-\eta(i+1)\Delta}}{(i+1)\eta} e^{-\eta(i+1)(k-1)\Delta}.
\end{aligned}
$$

Considering the ratio of the first term of the summation and $a\Delta\frac{1}{\eta}e^{-\eta(k-1)\Delta}\left[1-e^{-\eta\Delta}\right]$ yields

$$
\frac{1}{\Delta} \left(\frac{\sigma^2}{2\eta} \right) \frac{1}{\eta} \left[e^{\eta\Delta} - e^{-\eta\Delta} \right] e^{-\eta(k-l)\Delta}
$$

which is less than one for sufficiently large $\eta (k-l)$.

The variance expression is derived similarly. ■